Caracterización molecular y funcional de Small-Ubiquitin-related MOdifier en *Arabidopsis thaliana*

Laura Castaño Miquel

Universitat de Barcelona
Facultat de Farmacia
Departamento de Bioquímica y biología molecular

Centre de Recerca en Agrigenòmica (CRAG)

Caracterización molecular y funcional de Small-Ubiquitin-related MOdifier en *Arabidopsis thaliana*

Laura Castaño Miquel

Barcelona 2014

Universitat de Barcelona
Facultat de Farmacia
Departamento de Bioquímica y biología molecular

Programa de doctorato de Biotecnología

Caracterización molecular y funcional de Small-Ubiquitin-related MOdifier en *Arabidopsis thaliana*

Memòria presentada per Laura Castaño Miquel per optar al títol de doctor per la Universitat de Barcelona

La directora de tesi	**La doctoranda**	**La tutora**
Dra. L.Maria Lois Rojas	Laura Castaño Miquel	Dra.Josefa Badia

Barcelona 2014

Tenim a penes
el que tenim i prou: l'espai d'història
concreta que ens pertoca i un minúscul
territori per viure-la. Posem-nos
dempeus altra vegada i que se senti
la veu de tots, solemnement i clara.
Cridem qui som i que tothom ho escolti.
I, en acabat, que cadascú es vesteixi
com bonament li plagui, i via fora,
que tot està per fer i tot és possible.

Miquel Martí i Pol, 1982

ÍNDICE

ABREVIATURAS

ABA: ácido abscísico

ABI5: factor de transcripción aba insentive 5

ECFP: enchanced cian fluorescent protein

EIX: ethylene-induced xylanase

ESD4: early in short days 4

ET: etileno

EYFP: enchanced yellow fluorescent protein

FLC: flowering locus C

FLD: flowering locus D

HPY2: High Ploidy 2

HR: respuesta hipersensitiva

HsfA2: Arabidopsis heat shock factor A2

HSP: heat shock proteins

ICE1: inducer of CBF Expresion 1

JA: jasmónico

Mms21: methyl methanesulfonate-sensitive

NDSM: negatively charged amino acid-dependent SUMO motif

NES*:* Nuclear export signal

NLS: Nuclear localitzation signal

NPC: nuclear pore complex

PAMPs: pathogen-associated molecular patterns

PIAS: protein inhibitor of activated STAT

PCD: programa de muerte celular

PDSM: phosphorylation-dependent SUMO motif

PEG: polietilenglicol

PEP: posterior error probability

PR: pathogen response

PTMs: modificaciones post-traduccionales

ROS: especies reactivas de oxígeno

SA: ácido salicílico

SAE1: SUMO-activating enzyme subunidad 1

SAE2: SUMO-activating enzyme subunidad 2

SAR: respuesta sistemática adquirida

SAP: scaffold attachment factor (SAF)-A/B, acinus

SCE: SUMO-conjugating enzyme

SIZ1: PIAS SUMO E3 ligasa

SIM: SUMO interacting motifs

SIZ: SUMO E3 ligasa

SUMO: Small Ubiquitin-related MOdifier

TAIR: Arabidopsis information resource

TDG: thymine DNA glycosylase

UBLs: Ubiquitin like proteins

UFD: ubiquitin fold domain

ULP: Ubiquitin like proteases

INTRODUCCIÓN GENERAL

La complejidad generada por las modificaciones post-traduccionales

La vida y desarrollo celular se basa en los mecanismos de transcripción y traducción del ADN y el ARN. Estos mecanismos, ampliamente estudiados y regulados, finalizan con la síntesis y plegamiento de las proteínas, los actores finales en la célula. Cada gen contiene la información genética condensada de alta complejidad y su destino final es diverso y va en función de los requerimientos y/o desregulaciones celulares. Por tanto, a partir de un genoma de tamaño relativamente reducido, la célula es capaz de obtener variedad de funciones y regulaciones con un amplio proteoma. Uno de los principales mecanismos generadores de la amplia variedad del proteoma, son las modificaciones post-traduccionales (PTMs, *p*ost-*t*ranslational *m*odification*s*). Dichas modificaciones llegan a aumentar de manera drástica, hasta 50-100 veces, la diversidad del proteoma. Para ello, en la célula existen cientos de tipos de PTMs aunque, en general, se basan en la rotura de enlaces covalentes o proteólisis de la proteína modificada o en la formación de nuevos enlaces covalentes mediante la formación de puentes di-sulfuro, unión de pequeñas moléculas a un aminoácido específico o unión de otras proteínas mediante un enlace peptídico.

Las diferentes PTMs afectan la estructura de la proteína regulando la localización subcelular, la estabilidad, la actividad e interacciones proteína-proteína de la proteína diana modificada. Son un mecanismo rápido de regulación proteica en respuesta a señales intra y extracelulares. Es clave entender cómo, cuándo y por qué se dan para comprender la regulación y funcionamiento de cada una de las proteínas que conforman el proteoma. Además las modificaciones post-traduccionales son esenciales para el correcto funcionamiento y equilibrio de las proteínas dentro de la célula; defectos de éstas están directamente asociados a problemas de desarrollo y enfermedades. La adición de pequeñas proteínas o péptidos como la ubiquitina o proteínas de la familia de la ubiquitina se habían considerado específicas de sistemas eucarióticos, pero recientemente se han identificado sistemas análogos en procariotas (Maupin-Furlow, 2014).

La Ubiquitina y SUMO

En la década de los 80s se identificó en eucariotas la ubiquitina, la primera proteína con actividad de modificador post-traduccional. Llamándose así por su localización ubicua (en diferentes tejidos y especies). La ubiquitina está altamente conservada entre las especies; consta de 76 aminoácidos y tiene un peso molecular de unos 8,5kDa (Goldstein et al., 1975; SCHLESINGER and GOLDSTEIN, 1975). Esta modificación es reversible y es el resultado de la acción de una cascada enzimática conservada, y bioquímicamente bien caracterizada: la ubiquitinización. La unión de la ubiquitina, en general, determina la degradación del sustrato vía proteosoma, regulando la vida media de las proteínas (Hershko et al., 2000). Desde su descubrimiento, una familia entera de pequeñas proteínas relacionada con la ubiquitina ha sido descrita y nombrada como "Ubiquitin-like proteins" (UBLs).

SUMO, Small-Ubiquitin related MOdifer, es una proteína de la familia de la ubiquitina que fue descubierta en 1996. Fue identificada en células mamíferas cuando se estudiaba el funcionamiento del transporte nuclear activo mediado por la proteína RanGAP. Mediante el uso de anticuerpos específicos que reconocían RanGAP se identificaron dos formas de esta proteína que tan sólo diferían de unos 20kDa. Las secuencias peptídicas obtenidas confirmaron que se trataba de RanGAP más otras secuencias peptídicas no identificadas hasta el momento. Estas secuencias peptídicas presentaban un 18% de similitud con la Ubiquitina (Matunis et al., 1996). El experimento clave fue la incubación de la proteína recombinante RanGAP con el extracto celular de las células HeLa; al analizar el producto obtenido se apreció el cambio de masa molecular de RanGAP, indicando que esta proteína es modificada post-traduccionalmente por otra que fue bautizada con el nombre de SUMO (Mahajan et al., 1997). Así fue como se dio a conocer una nueva proteína, SUMO, que comparte estructura y ciclo similar al de la ubiquitina aunque difieren de su función. La ubiquitina mayoritariamente está relacionada con la vida media de las proteínas, regulando el estado celular, mientras que SUMO se ha asociado a diferentes procesos como el mantenimiento de la cromatina, la reparación del DNA, el transporte subcelular, la regulación transcripcional y la señalización celular (Hay, 2005).

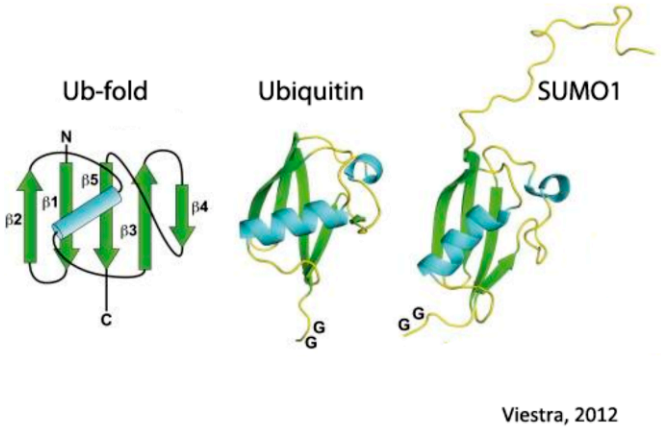


Figura 1. ***Diagrama de la estructura tridimensional de la ubiquitina y SUMO.*** A la izquierda, se muestra la estructura esquemática del dominio Ub-fold, ββαβββ, en el cual las hojas β se disponen alrededor de la hélice α orientada diagonalmente. A continuación, se muestra la estructura de la ubiquitina y SUMO humanas.

Aunque, solamente presentan un 18% de similitud de secuencia, SUMO y ubiquitina presentan una estructura terciaria bien conservada denominada Ub-fold, caracterizada por un plegamiento de hojas β alrededor de una hélix α (ββαβββ) (Bayer et al., 1998; Vijay-Kumar et al., 1986). Este dominio estructural, Ub-fold, esta altamente conservado entre las UBLs, de tal manera que las estructuras tridimensionales de SUMO y la ubiquitina son superponibles (Figura 1). SUMO a diferencia de la ubiquitina consta de unos 20 aminoácidos adicionales en el extremo amino terminal que dan lugar a un extremo N-terminal flexible. Ambas proteínas constan de un núcleo hidrófobo, mientras que la carga superficial de la enzima difiere significativamente en ambas, sugiriendo una especificidad del sistema (Bayer et al., 1998). Esta modificación covalente tiene lugar entre la glicina del extremo C-terminal de las UBLs y un residuo de lisina del sustrato; dando lugar a un enlace isopeptídico. La conjugación de las UBLs se da mediante una cascada enzimática similar a la de la ubiquitina, aunque cada uno de los enzimas que participan son únicos para cada modificador (Kerscher et al., 2006).

Ciclo de SUMOilación

La conjugación de SUMO al sustrato se da mediante una cascada enzimática denominada sumoilación, similar a la de la ubiquitinización. Esta vía consta de tres etapas donde participan tres enzimas: la enzima activadora de SUMO (E1) que cataliza la primera etapa de activación, la enzima conjugadora de SUMO (E2) responsable de la transferencia de de SUMO al sustrato en un reacción facilitada por la enzima ligasa E3.

SUMO es sintetizado como precursor y es necesaria su maduración para que sea activo. Son las proteasas específicas de SUMO (ULP, Ubiquitin Like Proteases) las que llevan a cabo su maduración, dejando expuesto el motivo di-glicina del extremo C-terminal, necesario para la conjugación de SUMO al sustrato. Seguidamente SUMO es activado mediante una reacción dependiente de ATP catalizada por la enzima activadora de SUMO, E1, heterodímero formado por dos subunidades, la subunidad grande SAE2 y la subunidad pequeña SAE1. En esta etapa la enzima activadora cataliza la hidrólisis del ATP para dar lugar al SUMO adenilado (SUMO-AMP). Seguidamente el AMP es liberado del SUMO adenilado y se genera un enlace tioéster entre el grupo carboxilo de la glicina de SUMO y el grupo sulfhidrilo de la cisteína activa del E1, situado en la subunidad grande SAE2 (SUMO-SAE2/SAE1) (Johnson, 2004). En la etapa de conjugación, el SUMO activado es transferido desde SAE2 a una cisteína de la enzima conjugadora, E2, mediante una reacción de transesterificación (SUMO-E2). Finalmente, en el paso de ligación SUMO es unido a la proteína diana mediante un enlace isopeptídico entre la glicina del extremo C-terminal de SUMO y la lisina del sustrato del motivo consenso, ΨKXE/D (Ψ, residuo alifático; L, lisina; X, cualquiera aminoácido; E/D glutamato o aspartato) (Rodriguez et al., 2001). En la etapa de ligación participan las enzimas E3 ligasas, facilitando la transferencia de SUMO desde el E2 al sustrato *in vitro e in vivo* (Johnson and Gupta, 2000; Pichler et al., 2001); aunque se ha demostrado que las proteínas recombinantes E1, E2 y SUMO son suficientes para la conjugación *in vitro*, característica conservada evolutivamente (Bernier-Villamor et al., 2001; Castaño Miquel et al., 2011).

La modificación por SUMO es reversible y la desumoliación esta mediada por la mismas proteasas (ULPs) que intervienen en la maduración de SUMO. En esta etapa, SUMO es reciclado y vuelve a actuar como modificador post-traduccional (Figura 2).

Además de los 4 aminoácidos canónicos que conforman el motivo consenso de sumoilación se han identificado otros: el motivo consenso invertido, el motivo hidrófobo de sumoilación, el motivo consenso de SUMO dependiente de fosforilación (PDSM, phosphorylation-dependent SUMO motif) y el motivo consenso de SUMO dependiente de carga negativa (NDSM, negatively charged amino acid-dependent SUMO motif) (Gareau and Lima, 2010; Vertegaal, 2011). La mayoría de los motivos de sumoilación se encuentran expuestos en la superficie o en loops extendidos y desorganizados del sustrato, adoptando una conformación extendida que facilita su reconocimiento por la enzima conjugadora y ligasa de SUMO.

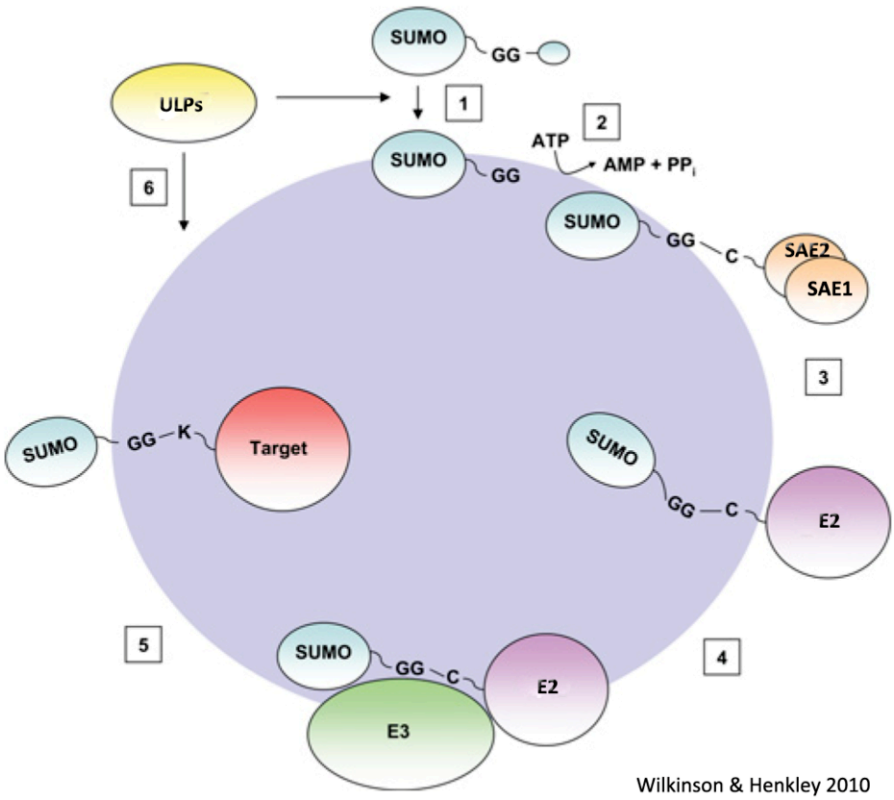


Figura 2. ***Representación esquemática del ciclo de SUMOilación.*** *(1)* SUMO es madurado por proteasas específicas que dejan expuesto el dominio di-glicina del extremo C-terminal. *(2)* Posteriormente, SUMO es activado en una reacción dependiente de ATP y unido mediante un enlace tioéster a la enzima activadora de SUMO. *(3)* Mediante una reacción de transesterificación SUMO es transferido a la enzima conjugadora de SUMO. *(4,5)* Finalmente SUMO es unido a la proteína diana reconocida por el E2 y E3. *(6)* La unión de SUMO es reversible y SUMO es desconjugado del sustrato por las mismas proteasas encargadas de su maduración, y puede ser reciclado entrando otra vez en la vía de conjugación de SUMO.

La modificación por SUMO puede darse como monómero o mediante la formación de cadenas de SUMO. La función biológica de las cadenas de SUMO no está clara. Estudios iniciales en levadura indicaban que no son esenciales (Bylebyl et al., 2003). Aunque, más recientemente, en mamíferos se ha descrito que la formación de cadenas de SUMO promueve la ubiquitinización del sustrato modificado señalizando para su degradación vía proteosoma (Tatham et al., 2008). Hasta hace relativamente poco tiempo, sólo unos pocos substratos de SUMO habían sido identificados. Los motivos principales recaen en el hecho que tan sólo alrededor de un 5% de la proteína diana es modificada por SUMO *in vivo*. Además, tras la lisis celular la mayoría de sustratos son susceptibles de desconjugación debido a la acción de las isopeptidasas. A pesar de estos desafíos, es crucial la identificación de sustratos de SUMO para comprender la función de esta modificación y su relevancia biológica.

Consecuencias moleculares de la sumoilación

Aunque las consecuencias funcionales de la unión de SUMO al sustrato son difíciles de predecir, de manera general podemos decir que su unión ocasiona un cambio en la superficie de interacción del sustrato y otras macromoléculas. Se han descrito tres principales consecuencias moleculares del sustrato ocasionadas por SUMO:

A. Interfiere en la capacidad de interacción con otras enzimas (figura 3A). Un ejemplo, es la modificación por SUMO en la lisina 14 de la enzima conjugadora de la ubiquitina E2-25K, que resulta en la inhibición de la interacción entre la enzima activadora y conjugadora de la ubiquitina; interfiriendo en el ciclo del sistema de la ubiquitina y por tanto disminuyendo la capacidad de ubiquitinización de los sustratos (Pichler et al., 2005).
B. Genera un nuevo sitio de unión capaz de reclutar otras macromoléculas, mediante la creación de una nueva superficie de interacción o interacciones no covalentes con SUMO mediadas por los motivos SIMs (SUMO interacting motifs) (figura 3B). Los SIMs interactúan no covalentemente con SUMO y se caracterizan por un núcleo hidrófobo flanqueado por aminoácidos ácidos (mayoritariamente serinas, Ser) (Song et al., 2004). De hecho, la sumoilación de RanGAP promueve la interacción con RanBP2; RanBP2 consta de dos motivos SIMs en el extremo carboxi terminal que promueven la unión al complejo SUMO-RanGAP y la relocalización de este complejo al poro nuclear (NPC, Nuclear Pore Complex) (Werner et al., 2012).
C. Cambio conformacional de la proteína diana (figura 3C). Suponiendo que el sustrato contenga un motivo SIM, se puede dar una interacción intramolecular entre SUMO y el motivo SIM de la proteína modificada resultando en un cambio de conformación de la proteína (Kerscher et al., 2006). Este es el caso de la enzima TDG (Thymine DNA Glycosylase), que se une al DNA con el fin de reparar errores. La sumoilación de TDG da lugar a un cambio conformacional de la proteína mediante la interacción del SUMO unido covalentemente y el motivo SIM de TDG, que permiten la escisión de TDG del DNA para que pueda volver al ciclo de reparación de DNA (Hardeland et al., 2002).

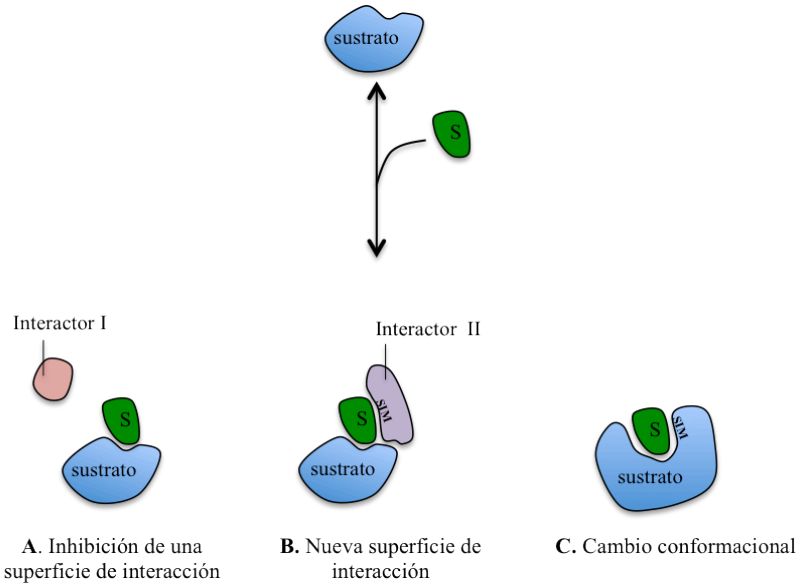


Figura3. ***Consecuencias moleculares de la SUMOilación. (a)*** La modificación por SUMO interfiere con la interacción de la proteína diana con otras proteínas. ***(b)*** El resultado de la SUMOilación da lugar a una nueva superficie de interacción capaz de interaccionar con otras proteínas, algunas de ellas se dan entre el motivo SIM del interactor y el complejo SUMO-substrato. ***(c)*** La sumoilación puede dar lugar a un cambio conformacional del sustrato. (S:SUMO; SIM: SUMO interacting motif).

Componentes de la maquinaria de conjugación en *Arabidopsis Thaliana*

Con el fin de entender la regulación y la relevancia biológica de la SUMOilación en organismos vegetales, se ha escogido como modelo de estudio *Arabidopsis thaliana*. La secuenciación del genoma de *Arabidopsis* ha permitido identificar los genes que codifican para cada uno de los componentes del sistema de SUMOilación, confirmando que este sistema se encuentra conservado en *Arabidopsis*. El estudio bioquímico de las enzimas implicadas en esta vía y estudios de ganancia o pérdida de función, han elucidado mecanismos de regulación de esta modificación y algunas de sus implicaciones biológicas. En la tabla 1 se encuentra un resumen de las enzimas que se encuentran expresadas en el sistema de SUMOilación de *Arabidopsis thaliana*.

Tabla 1

Componentes de la maquinaria de SUMO en *Arabidopsis thaliana*

				Fenotipo	
Proteína	Gen	Gen	Localización	Mutante	Doble mutante
SUMO	*At*SUMO1	At4g26840	Citoplasma/nuclear	Normal	*atsumo1/atsumo2*; Letal
	*At*SUMO2	At5g55160	Citoplasma/nuclear	Normal	*sum1-1/amiR-SUM2*; pleiotrópico
	*At*SUMO3	At5g55170	-	Normal	
	*At*SUMO5	At2g32765	-	-	
E1	*At*SAE2	At2g21470	-	Letal	
	*At*SAE1a	At4g24940	-	Normal	
	*At*SAE1b	At5g50580	-	-	
E2	AtSCE1	At3g57870	Citoplasma/nuclear	Letal	
E3	*At*SIZ1	At5g60410	Nuclear	Pleiotrópico	*hpy2-1/siz1-2;* Letal
	AtHYP2	At3g15150	Nuclear	Pleiotrópico	
ULP	*At*ULP1a	At3g06910	citoplasma	Normal	atulp1a/atesd4; pleiotropic
	*At*ULP1c	At1g10570	Nuclear	Normal	*atulp1c/atulpc1d;* sensibilidad salina
	*At*ULP1d	At1g60220	Nuclear	Normal	
	*At*ESD4	At4g75880	Poro nuclear	Enano/Pleiotrópico	

SUMO: small ubiquitin-related modifier; E1: dos subunidades codifican para la enzima activadora de SUMO SAE2 y SAE1; E2: un solo gen codifica para la enzima conjudora, SCE1; E3: hasta el momento dos tipos de E3 ligasas han sido identificadas, SIZ1 y HYP2; ULPs: cuatro enzimas codifican para las ubiquitin-like proteasas. Se muestra ell fenotipo de los mutantes y dobles mutantes analizados hasta el momento. Tabla adaptada de Lois, 2010.

SUMO

En levadura, moscas y nemátodos existe tan sólo un gen que codifica por SUMO, mientras que en vertebrados y plantas hay una familia entera de genes que codifican para SUMO. En el genoma de *A.thaliana* hay ocho genes que codifican para las diferentes isoformas de SUMO, aunque, solamente se han detectado niveles de expresión significativos para *At*SUMO1, 2, 3 y 5. *At*SUMO1 y *At*SUMO2 están altamente conservados, presentan un 89% de identidad de secuencia; mientras que *At*SUMO3 y *At*SUMO5 presentan un 42% y 30% de identidad de secuencia con *At*SUMO1, respectivamente (Kurepa, 2002). Alineamientos de secuencia con sus ortólogos humanos, levadura y drosophila muestran que la región más conservada se encuentra en la región Ub-fold y la más divergente corresponde al N-terminal de SUMO (Kurepa, 2002). La expresión *in vivo* de *At*SUMO1/2 y *At*SUMO3 se ha evidenciado mediante anticuerpos que reconocen ambas proteínas (Kurepa, 2002; Lois, 2003), mientras que *At*SUMO5 tan sólo se ha podido detectar *in vivo* en plantas sobrexpresoras de *At*SUMO5 (Budhiraja et al., 2009).

El estudio de los parálogos de SUMO en mamíferos y plantas ha puesto de manifiesto que han adquirido distintas propiedades moleculares y funcionales. En plantas, *At*SUMO1 y 2 son capaces de formar cadenas de SUMO mientras que *At*SUMO3 y 5 no constan de motivo de SUMOilación (Colby et al., 2006). Asimismo, las proteasas de SUMO presentan una baja capacidad isopeptidasa respecto a *At*SUMO3 y una alta capacidad de desconjugación de conjugados de *At*SUMO1 y 2 (Chosed et al., 2006). Mediante análisis genéticos se ha determinado el papel esencial de *At*SUMO1 y 2. El doble mutante *atsumo1/atsumo2* es letal,

causando el arresto del desarrollo embrionario en las primeras etapas, indicando que *At*SUMO3 y 5 no son capaces de complementar la función biológica de *At*SUMO1 y 2 (Saracco et al., 2007). Además, se ha observado una acumulación diferencial de conjugados de los parálogos de SUMO en respuesta a diferentes tipos de estrés (estrés térmico, peróxido de hidrogeno y etanol): se da una rápido aumento de conjugados de *At*SUMO1 y 2 mientras que no se observa con *At*SUMO3 (Kurepa, 2002). Esta especialización de los parálogos de SUMO en Arabidopsis se ve reforzada por una diferente localización y abundancia de las proteínas; *At*SUMO1 y 2 se encuentran altamente expresadas, mientras que la expresión de *At*SUMO3 es generalmente baja. Además, *At*SUMO1 se expresa en los diferentes tejidos y estadios de desarrollo de la planta, sugiriendo un papel más genérico de esta isoforma, mientras que *At*SUMO2 presenta una localización más específica, cómo en el tejido vascular de la hoja (van den Burg et al., 2010).

La enzima Activadora de SUMO, E1

La enzima activadora de SUMO es la encargada de llevar a cabo la reacción de adenilación, donde se da la activación de SUMO y su posterior transferencia a la enzima conjugadora. La cristalización de la enzima activadora de la ubiquitina, SUMO y Nedd8 ha revelado que existe un modelo estructural conservado para la enzima E1(Lee and Schindelin, 2008; Lois and Lima, 2005; Walden et al., 2003). En contraste con la enzima activadora de la ubiquitina, la de SUMO es un heterodímero que conserva los dominios estructurales y funcionales de la enzima activadora de la ubiquitina. En humanos, *S.cerevisiae* y *Arabidopsis thaliana* está formada por dos subunidades: la subunidad grande (SAE2) y la subunidad pequeña (SAE1).

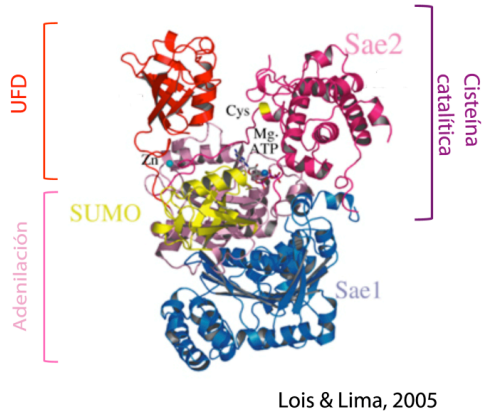


Figura 4. ***Estructura de la enzima activadora de SUMO.*** Diagrama del complejo SAE1/SAE2-SUMO-Mg·ATP. Los dominios de SAE2 se encuentran representados en diferentes colores: rojo para el dominio UFD, magenta para el dominio de la cisteína catalítica y rosa pálido para el dominio de adenilación. La cisteína catalítica activa esta resaltada en amarillo. La subunidad SAE1 se encuentra representada en azul y finalmente SUMO en amarillo.

La subunidad grande consta de 3 dominios funcionales: el dominio de adenilación, el de la cisteína catalítica y el ubiquitin-fold domain (UFD) (figura 4). Todos ellos juegan un papel esencial en la primera etapa del ciclo de SUMO. El dominio de adenilación es el responsable del reconocimiento y adenilación del C-terminal de SUMO. Este dominio se encuentra flanqueado por el de la cisteína catalítica y UFD (Figura 4). El dominio de la cisteína catalítica es necesario para la formación del enlace tioéster E1-SUMO, que va acompañado de un cambio conformacional del E1 con el fin de aproximar el SUMO adenilado a la cisteína catalítica (Olsen et al., 2010). La transferencia de SUMO desde la enzima activadora a la conjugadora, se da por un previo reconocimiento del E2 por el E1, mediado por el dominio UFD; este dominio tal y como indica su nombre tiene una estructura tridimensional parecida a la ubiquitina, y es imprescindible para el reconocimiento del E2 l(Lois and Lima, 2005). Estudios estructurales y bioquímicos han determinado que los aminoácidos implicados en la interacción entre SAE2^{ufd}-E2 están altamente conservados en humanos y levadura (Wang et al., 2010). En este reconocimiento participa también el dominio de la cisteína catalítica, acercando los dos residuos de cisteína del E1 y E2, proceso esencial para la transferencia de SUMO desde el E1 al E2 (Wang et al., 2007).

En *Arabidopsis thaliana,* existen dos formas de la enzima activadora, E1a y E1b, que difieren en la composición de la subunidad pequeña del E1, *At*SAE1a y *At*SAE1b, y presentan un 81% de similitud de secuencia. SAE2 es esencial, la inactivación de la etapa de activación mediante mutantes de *atsae2* son letales, impidiendo el desarrollo del embrión (Saracco et al., 2007).

La enzima conjugadora de SUMO, E2

La enzima conjugadora de SUMO tiene un papel central en el ciclo de sumoilación, interactuando con todos los enzimas de la vía. Esta enzima es inicialmente reconocida por el dominio UFD y el dominio que contiene la cisteína catalítica de la E1, facilitando así la transferencia de SUMO desde el E1 al E2. Una vez formado el enlace tioéster entre la glicina del C-terminal de SUMO y la cisteína catalítica del E2, éste cataliza la unión de SUMO al sustrato directamente o mediante la participación de las E3 ligasas.

Esta enzima confiere especificidad de sustrato reconociendo mediante interacciones no covalentes el motivo consenso de sumoilación del sustrato, acercando la proteína diana al SUMO unido covalentemente al E2 (Lin et al., 2002). La superficie responsable de este reconocimiento se encuentra adyacente a la cisteína catalítica del E2 (Gareau and Lima, 2010). Además de formar el intermedio tioéster con SUMO, la enzima conjugadora de

SUMO es capaz de formar un complejo con SUMO mediante interacciones no covalentes. En mamíferos, la interacción no covalente entre E2 y SUMO promueve la formación de cadenas de SUMO (Knipscheer et al., 2007).

A diferencia del sistema de la ubiquitina, en la que se han identificado diferentes enzimas E2 tan sólo existe una para SUMO en mamíferos, levadura y plantas. Sugiriendo que la enzima conjugadora de SUMO juega un papel en la especificidad de sustrato junto con las E3 ligasas (Gareau and Lima, 2010). El genoma de *A.thaliana* contiene un pseudogén y un gen activo que codifican por la enzima conjugadora de SUMO, *At*SCE1b y *At*SCE1a respectivamente. Existe tan sólo una isoforma del E2 codificada por *At*SCE1a y la inactivación de este gen es letal (Saracco et al., 2007). Se ha detectado la expresión del E2 en casi todos los tejidos de *Arabidopsis,* y muestra una colocalización nuclear junto con *At*SUMO1/2 (Lois, 2003).

Las enzimas ligasas de SUMO, E3

Estas enzimas cumplen tres propiedades: interaccionan con la enzima conjugadora de SUMO, con el sustrato y facilitan la transferencia de SUMO desde el E2 a la proteína diana. Las SUMO E3 ligasas identificadas en eucariotas se caracterizan por la presencia del dominio SP-RING, parecido al dominio RING que contienen la mayoría de las ligasas E3 de la ubiquitina (Johnson, 2004). Este dominio es esencial para la interacción con la enzima conjugadora de SUMO y tiene actividad ligasa. Las PIAS (Protein Inhibitor of Activated STAT) son las SUMO E3 ligasas más caracterizadas y abundantes. Cinco dominios diferentes han sido identificados: SAP, PINIT, SP-RING, SIM y el dominio C-terminal rico en serina y threonina. El dominio SAP (Scaffold Attachment factor (SAF)-A/B, acinus, PIAS) del extremo amino terminal es requerido para la unión al DNA. El dominio PINIT es requerido para la localización nuclear de la proteína. Y finalmente, el dominio SIM que permite la interacción no covalente con SUMO (Rytinki et al., 2009; Sharrocks, 2006). La clase de SUMO E3 ligasas que no contienen el dominio PINIT y SAP se denominan Mms21 (Methyl Methanesulfonate-Sensitive).

Hasta el momento se han identificado dos SUMO E3 ligasas en *Arabidopsis* mientras que hasta tres o 15 han sido descritas en el sistema de levadura y mamíferos, respectivamente (Miura and Hasegawa, 2010; Wilkinson and Henley, 2010). La primera identificada fue SIZ1 (PIAS SUMO E3 ligasa) perteneciente a la familia de las PIAS. *At*SIZ1 conserva los cuatro dominios funcionales de las PIAS y contiene además el dominio PHD (Plant HomeoDomain), el cual interacciona con SCE y tiene actividad ligasa (Garcia-Dominguez et al., 2008). La segunda SUMO E3 ligasa fue identificada por dos grupos independientes y

nombrada MMS21 y HPY2 (*M*ethyl *M*ethanesulfonate-Sensitivity protein *21*, *H*igh *P*loidy *2*), respectivamente (Huang et al., 2009; Ishida et al., 2009) a la que nos referiremos como HPY2. Los mutantes de las SUMO E3 ligasas no son letales y han permitido elucidar el papel de la sumoilación en las plantas. Mutantes de pérdida de función de SIZ1 y HPY2 causan un fenotipo enano, el fenotipo *athpy2-1* no es complementado con la expresión de SIZ1 y el doble mutante es letal. Estos resultados sugieren que SIZ1 y HYP2 no son redundantes y funcionan independientemente, la combinación de ambas es esencial para un correcto funcionamiento y desarrollo de la planta (Ishida et al., 2011).

Proteasas específicas de SUMO, ULPs

La balanza celular entre las concentraciones de SUMO libre y conjugado esta regulado por las ULPs. Estas enzimas tienen actividad peptidasa, requisito para la maduración de SUMO, y actividad isopeptidasa para la desconjugación de SUMO del sustrato. Las SUMO proteasas consisten en un extremo amino terminal variable y el extremo carboxilo terminal conservado con actividad proteasa.

Las proteasas de SUMO forman la familia más extensa de los diferentes componentes de la maquinaria de conjugación. Aunque, en *Arabidopsis thaliana* sólo se han detectado niveles de expresión significativos para cuatro isoformas: *At*ULPa, *At*ULP1c, *At*ULPd y *At*ESD4 (Early in Short Days 4) (Colby et al., 2006; Kurepa, 2002; Murtas, 2003). Todas ellas muestran actividad peptidasa o isopeptidasa frente *At*SUMO1 y 2, mientras que tan sólo *At*ULP1a tiene actividad peptidasa frente *At*SUMO3 y ninguna de ellas para *At*SUMO5 (Chosed et al., 2006; Colby et al., 2006). No se ha podido detectar actividad peptidasa endógena para *At*SUMO3/5 en *Arabidopsis*, sugiriendo que *At*SUMO3/5 deben ser madurados en un nivel bajo o en circunstancias específicas (Lois, 2010). *At*ULP1a presenta localización citoplasmática, mientras que *At*ULP1c, *At*ULP1d y *At*ESD4 muestran localización nuclear (Conti et al., 2008; Murtas, 2003).

Estudios genéticos han demostrado que *At*ULP1c y *At*ULP1d tienen funciones redundantes; la pérdida de función de una de estas dos isoformas no resulta en ningún fenotipo significante mientras que el doble mutante *atulp1d/atulp1c* presenta floración temprana y sensibilidad salina (Conti et al., 2008). El mutante *At*ULP1a tampoco presenta fenotipo, a diferencia de su homólogo más cercano ESD4 que presenta un fenotipo mucho más dramático (Hermkes et al., 2010). Los mutantes *esd4* presentan unos niveles reducidos de SUMO libre y un aumento de conjugados. La sobreexpresión de la forma endógena de SUMO maduro en el mutante *esd4* acentúa los defectos de desarrollo que confiere la mutación *esd4*, poniendo de

manifiesto la importancia de la desconjugación de SUMO en el desarrollo de las plantas (Murtas, 2003).

Implicaciones biológicas de SUMO en las plantas

La unión de SUMO a las proteínas dianas da lugar a un cambio de actividad del sustrato, generando o bloqueando superficies de interacción de la proteína diana. El estudio de los efectos biológicos de la modificación post-traduccional mediada por SUMO ha elucidado diferentes funciones biológicas reguladas por SUMO en las plantas. Para ello ha sido crucial el estudio de mutantes de pérdida y ganancia de función de los diferentes componentes de la maquinaria de conjugación de SUMO y la identificación de sustratos de SUMO.

Estrés abiótico

La primera evidencia fue observada por Kurepa y colaboradores (Kurepa, 2002), los cuales descubrieron un aumento de conjugados de SUMO frente varios tipos de estreses abióticos (estrés térmico, peróxido de hidrogeno y etanol). Se ha estudiado en detalle la respuesta al estrés térmico, plántulas de *Arabidopsis* expuestas a 37 grados muestran un rápido y dramático aumento de conjugados de SUMO1 y 2 donde se consume casi todo el SUMO libre (Kurepa, 2002). Esta modificación es reversible y al cabo de unas horas el SUMO libre es regenerado. La sobre-expresión de SUMO1 reduce la tolerancia adquirida, sugiriendo que la sumoilación actúa negativamente sobre la termotolerancia adquirida (Cohen-Peer et al., 2010). HsfA2, Arabidopsis heat shock factor A2, es un factor de transcripción que induce la expresión de genes de respuesta al estrés térmico, como las heat shock proteins (HSP) (Kotak et al., 2007). La SUMOilación de HsfA2 reduce su actividad, reduciendo la termotolerancia adquirida de la planta (Cohen-Peer et al., 2010) (figura 5A). Por el contrario, los mutantes *siz1-3* presentan sensibilidad a la exposición a temperaturas altas, indicando que SIZ1 regula positivamente la termotolerancia basal (Yoo et al., 2006). Es un claro ejemplo de la complejidad de la sumoilación, donde diferentes componentes de la maquinaria pueden actuar antagonistamente en respuesta a diferentes tipos de estrés térmico, indicando la importancia de la regulación del estado celular de los conjugados de SUMO y su dinamismo. Estudios proteómicos han identificado diferentes sustratos de SUMO frente a estrés térmico, la gran mayoría están relacionados en el mantenimiento de la homeostasis del RNA, sugiriendo el papel de SUMO como regulador del transcriptoma. Además, muchas de las proteínas modificadas por SUMO en respuesta a estrés térmico también son sumoiladas

en condiciones de estrés oxidativo y etanol, indicando que su modificación es crucial en respuesta a estrés (Miller et al., 2013).

Además del papel de SUMO frente al estrés térmico, también se ha descrito la importancia de la E3 ligasa de SUMO SIZ1 en el control de la adaptación de las plantas a bajas temperaturas. La SUMOilación dependiente de SIZ1 activa o estabiliza la proteína ICE1 (*Inducer of CBF Expresion 1*), que controla la expresión de los factores de transcripción CBF3/DREB1A (*C-repeat binding factor 3/ Dehydration Responsive Element binding factor 1a*) de respuesta al frío. La SUMOilación de ICE1 estabiliza la proteína bloqueando su ubiquitinización. Además, la modificación de ICE1 por SUMO regula negativamente MYB15, represor de CBF3/DREB1A, mediante su unión a elementos MYB del promotor (Miura et al., 2007). SIZ1 regula positivamente la respuesta a bajas temperaturas (figura 5B).

El estudio del mutante *siz1* ha elucidado el papel de SUMO frente diferentes estreses abióticos. Estos mutantes presentan un patrón de sumoilación alterado que resulta en un aumento generalizado de sensibilidad frente a estos tipos de estrés. Se ha descrito el papel de la E3 ligasa SIZ1 de SUMO en el control de la respuesta frente déficit de fósforo, estrés hídrico, en el mantenimiento de la homeostasis del cobre y el estrés salino (Catalá et al., 2007; Chen et al., 2011; Miura et al., 2005; 2007; 2012).

Estrés biótico

Uno de los objetivos de los patógenos cuando infectan a su huésped es la proteólisis de las proteínas huéspedes. Curiosamente, se han identificado varios factores virulentos con actividad cisteína proteasa que posen actividad peptidasa y isopeptidasa frente a SUMO, actuando como proteasas específicas de SUMO. Estos factores virulentos desregulan la balanza de SUMO libre y conjugado de la célula, comprometiendo el sistema de defensa del huésped (Hotson and Mudgett, 2004).

Por otro lado, uno de los mecanismos de defensa de la planta frente a la infección por patógenos es la respuesta sistemática adquirida (SAR), que va precedida por una acumulación de acido salicílico (SA) que activa la cascada de señalización del SA activando una serie de genes de respuesta al patógeno (PR genes). Los mutantes *siz1-3* presentan una mayor acumulación de SA, por tanto tienen constitutivamente activada la respuesta SAR y una mayor expresión de los genes PR en condiciones basales, mediante la expresión del factor de transcripción NPR1 requerido para la activación de la respuesta sistemática adquirida. Esta cascada promovida por el SA da lugar a una resistencia de estos mutantes frente la infección del patógeno *Pseudomonas syringae* (*Pst*DC3000) (Lee et al., 2006b). Además, una mayor

acumulación de SA en las plantas *siz1-3* resulta en un cierre de los estomas que reduce la invasión por patógenos y confiere tolerancia frente estrés hídrico (Miura et al., 2012). Consistente con estos resultados, los mutantes *atsum1-1/amiR-SUM2* presentan acumulación de SA y resistencia frente *Pst*DC3000 (van den Burg et al., 2010) (figura 5C).

Floración

La primera vinculación entre SUMO y el control de la floración, fue descrita con el mutante de la proteasa de SUMO *At*ESD4 (Murtas, 2003). Posteriormente también se ha descrito la floración temprana de los mutantes *siz1-3* (Jin et al., 2008)*;* volviendo a producirse una situación donde dos componentes que actúan antagonistamente, aumentando o disminuyendo los conjugados de SUMO tienen el mismo efecto final. Ambos mutantes causan una reducción de expresión y actividad del FLC (flowering locus C), factor de transcripción que regula negativamente la transición floral. El mecanismo molecular por el cual FLC está regulado aún ésta por elucidar. Se ha descrito que la expresión de este factor, FLC, es sensible a la estructura de la cromatina, la cual puede estar afectada por un aumento conjugados de SUMO (Novatchkova et al., 2004). Por otro lado, FLC es reprimido por el factor FLD, la sumoilación de FLD (flowering locus D) dependiente de SIZ1 da lugar a su inactivación causando una activación del FLC y una represión floral (Jin et al., 2008) (figura 5D).

Desarrollo

Los mutantes *atsae2, atsce y atsumo1/2* son letales, el embrión no se desarrolla correctamente; pudiéndose establecer un vinculo importante entre la sumoilación y el desarrollo del embrión (Saracco et al., 2007). Con el fin de elucidar el papel de SUMO durante el desarrollo reproductivo se ha estudiado el papel funcional de las E3 ligasas durante este estado de desarrollo. Se ha descrito que la E3 ligasa *At*SIZ1 regula la estabilidad y desarrollo del gametófito femenino, necesario para la correcta orientación del tubo polínico y posterior fertilización, mientras que la ligasa *At*HYP2 es necesaria para el correcto desarrollo del gametofito femenino y masculino (Ling et al., 2012; Liu et al., 2014). Por otro lado, *atsiz1 y athyp2* tienen un fenotipo enano, aunque viene dado por diferentes mecanismos en cada uno de los casos. En el caso de los mutantes de la E3 ligasa *At*SIZ1 este fenotipo se ha asociado a la acumulación de SA (Lee et al., 2006b); por el contrario la disminución de la proliferación celular en los mutantes de la E3 ligasa *At*HYP2 es independiente de SA (Ishida et al., 2011).

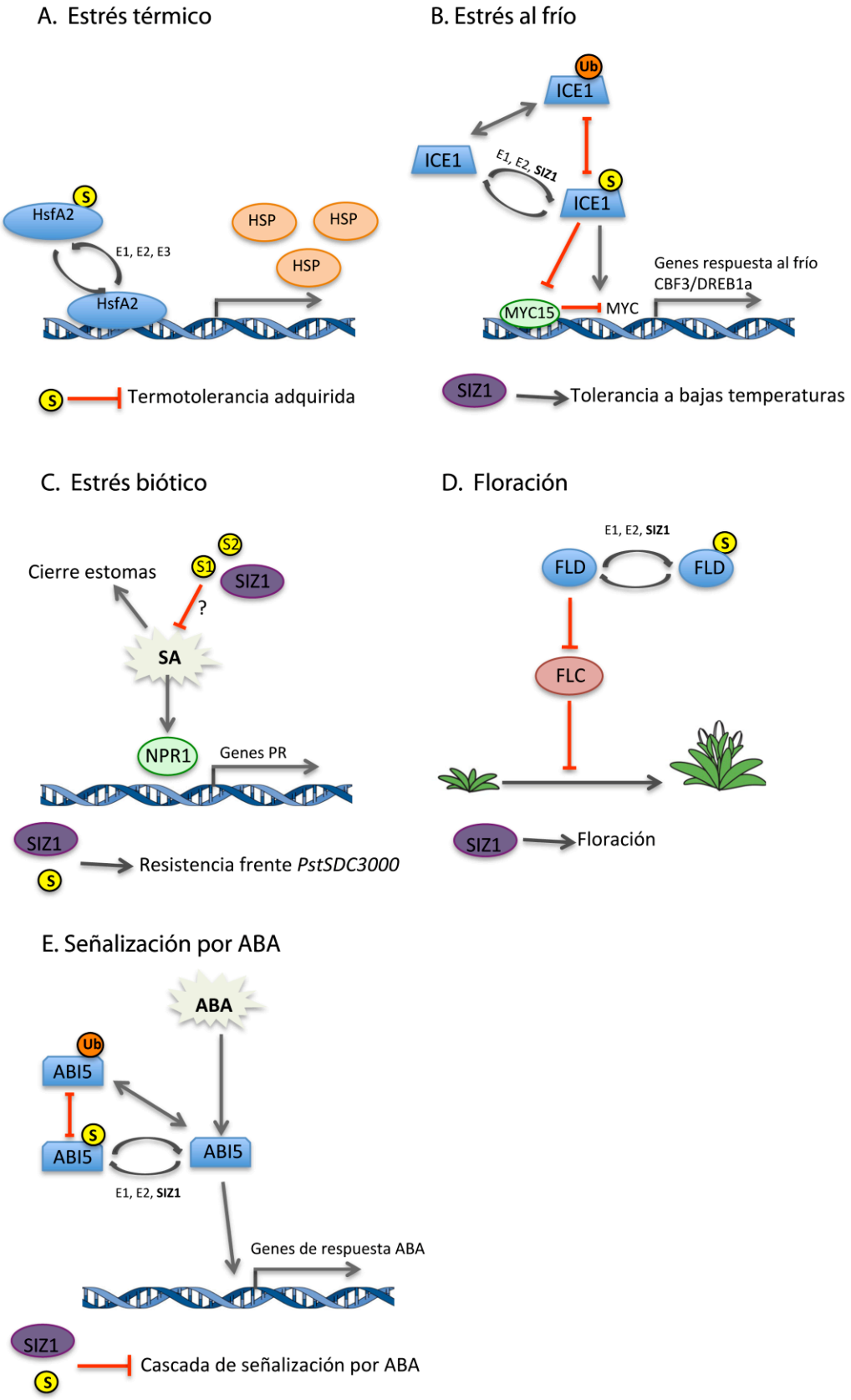


Figura 5. ***Efectos de la SUMOilación en las plantas***. ***(a)*** La SUMOilación regula negativamente la termotolerancia adquirida de la planta. El factor de transcripción HsfA2 es un elemente clave en la termotolerancia adquirida. La SUMOilación de este factor lo inactiva reduciendo la termotolerancia adquirida

de la planta. ***(b)*** La SUMOilación de ICE1 es dependiente de SIZ1. ICE1-SUMO antagoniza su ubiquitinización. Además inhibe MYB15, factor represor de CBF3/DREB1A. SIZ1 regula positivamente la respuesta a bajas temperaturas. ***(c)*** La invasión de patógenos en plantas genera un aumento de SA, la cascada de señalización resulta en la expresión de los genes PR. SIZ1 y SUMO son un regulador negativo de SA, los mutantes *siz1-3* y *atsum1-1/amiR-SUM2* presentan más resistencia frente *PstDC3000*. Además, el incremento de SA promueve el cierre de los estomas disminuyendo la entrada de patógenos. ***(d)*** Los mutantes *siz1-3* presentan floración temprana. La SUMOilación del factor FLD inhibe el represor floral de floración FLC. ***(e)*** La SUMOilación de ABI5 por SIZ1 regula negativamente la señalización por ABA. ABI5-SUMO es inactivo y es resistente a la degradación por proteosoma mediada por la ubiquitinización.

La respuesta a hormonas

En varias de las funciones biológicas comentadas anteriormente las hormonas juegan un papel importante. El ABA (ácido abscísico) es una fitohormona que regula numerosos procesos relacionados con el correcto desarrollo y adaptación de la planta a cambios ambientales. Plantas sobrexpresoras de *At*SUMO1 y 2 son insensibles al ABA, mientras que plantas con la expresión de *At*SCE1 reducida muestran un aumento de la sensibilidad al ABA (Lois, 2003). Todo ello indica que la SUMOilación regula negativamente las respuestas mediadas por el ABA. El estudio del efecto de ABA en los mutantes de la E3 ligasa *At*SIZ1, ha permitido identificar la sumoilación del factor ABI5 (ABA insenitive 5) dependiente de SIZ1. Además, como ya hemos descrito en otros casos, la SUMOilación de ABI5 evita su degradación vía proteosoma facilitada por AFP (ABI five binding protein) (Lopez-Molina et al., 2003). La sumoilación del factor ABI5 regula negativamente la señalización mediada por ABA en la célula (Miura et al., 2009), hormona que juega un papel esencial en respuesta a diversos estreses y desarrollo de la planta (figura 5E).

También se ha relacionado SUMO con el control del SA, regulando la respuesta del sistema inmune y desarrollo de la planta. Se ha observado una acumulación de SA en plantas sobrexpresoras de *At*SUMO1 y 2, en los mutantes *siz1* y *esd4;* sugiriendo que la homeóstasis de SUMO es vital para la modulación del SA en la célula (Lee et al., 2006b; van den Burg et al., 2010). Además, potencialmente SUMO se ha relacionado con la regulación del metabolismo de las auxinas, etileno, brasinoesteroides y citoquininas (Castro et al., 2012). Todas estas hormonas juegan un papel muy importante en la respuesta a diferentes estreses abióticos, bióticos y desarrollo de la planta. Entender la relación entre SUMO y las diferentes respuestas hormonales nos ayudará a elucidar el mecanismo molecular por el cual SUMO actúa.

Modulación del ciclo de SUMOilación

Con el fin de que la modificación por SUMO sea efectiva, el ciclo de sumoilación debe ser regulado. Los niveles de conjugados de SUMO en la célula están afectados por diferentes factores cómo ciclo celular, diferenciación, en respuesta a diferentes tipos de estrés y daño celular. Esta modulación de conjugados/desconjugados de SUMO se puede dar tanto por un aumento de actividad y viabilidad de los componentes de la maquinaria de SUMO cómo por un aumento de los niveles de estos mismos. Se ha descrito tanto una regulación de los componentes a nivel transcripcional como post-traduccional en diferentes organismos.

Regulación de SUMO a nivel transcripcional

Los niveles de los transcritos de algunos de los componentes de la maquinaria de SUMO en mamíferos se encuentran diferencialmente expresados en función del tejido, etapa de desarrollo, estrés ambiental y diferenciación celular (revisado por (Watts, 2013)). En plantas, concretamente en arroz, se ha descrito la regulación de la enzima conjugadora de SUMO, *Os*SCE1, en respuesta a diferentes estreses. Se ha observado un aumento del transcrito *Os*SCE1 en respuesta a altas temperaturas y una disminución del mRNA de *Os*SCE1 frente tratamiento con PEG (PoliEtilenGlicol, utilizado como agente osmótico para generar déficit de agua en las plantas). Los tránscritos de esta enzima están regulados en función de las condiciones ambientales, en arroz (Nigam et al., 2008). No obstante, en *Arabidopsis*, se observa una expresión a nivel de proteína diferencial de las enzimas de la maquinaria de SUMO durante el desarrollo de la planta que no correlaciona con cambios a nivel de mRNA; sugiriendo que debe estar regulado mayoritariamente a nivel post-traduccional (Silvia Manrique y Inês Teixeira, no publicado).

Relocalización y regulación a nivel post-traduccional

La localización de los componentes de la maquinaria de SUMO juega un papel importante en la regulación de la función del sistema de sumoilación. Aunque la gran mayoría de los sustratos de SUMO identificados presentan una localización nuclear, algunos de ellos están presentes en otros compartimientos celulares. Curiosamente, la mayoría de los componentes de la maquinaria de SUMO presentan una localización nuclear predominante. En este contexto, se ha descrito que la relocalización de las enzimas del sistema de SUMO es importante y es un mecanismo de regulación de la conjugación de SUMO. Esté es el caso de *Saccharomyces cerevisiae,* tiene dos isoformas que codifican para las proteasas específicas de SUMO, Ulp1 y Ulp2, y ambas muestran una localización diferente: Ulp1 esta anclada al

NPC y Ulp2 es nuclear. Estas proteasas tienen sustratos diferentes y esta especificidad de sustrato viene dada en gran parte por la localización de las enzimas. Teniendo en cuenta que formas truncadas de la isoforma Ulp1 muestran una localización nuclear de la enzima capaz de complementar el mutante *ulp2*, la especificidad de sustrato de Ulp1 parece estar determinada en parte por su localización (Li and Hochstrasser, 2003). Además, Ulp1 es reclutado al nucléolo de la célula en respuesta a la toxicidad celular inducida por alcohol dando lugar a un aumento de conjugados celulares. El secuestro de Ulp1 al nucléolo disminuye la desconjugación de SUMO, perturbando el equilibrio de conjugación-desconjugación resultando en un aumento de conjugados de SUMO en respuesta a este tipo de estrés (Sydorskyy et al., 2010). El enigma recae en cuál es el factor que regula esta relocalización de la enzima Ulp1, la cuál aun no ha sido resuelta. Se ha descrito también un cambio de localización celular de la E3 ligasa de levadura durante el ciclo celular dependiente de la fosforilación de la enzima (Johnson and Gupta, 2000). En algunos casos la regulación del transporte celular de las enzimas viene dada por modificaciones post-traduccionales. Éste es el caso de la enzima activadora de SUMO en humanos, que es retenida en el núcleo mediante la sumoilación de la subunidad grande de la enzima (Truong et al., 2012b). Además, se ha detectado la sumoilación de la enzima conjugadora y ligasa de SUMO en levadura y humanos, aunque no está establecida una relación con su localización subcellular. En el caso de la enzima conjugadora se ha postulado que su sumoilación regula la especificidad de sustrato, generando una nueva superficie de interacción con un aumento de afinidad por el sustrato (Knipscheer et al., 2008). En plantas, se ha identificado la sumoilación de SAE2, SIZ1 y SCE *in vivo*, aunque no se ha elucidado el papel biológico de estas modificaciones post-traduccionales (Elrouby and Coupland, 2010; Miller and Vierstra, 2010).

La sumoilación y la fosforilación no son las únicas PTMs de las que son sustrato las enzimas de la vía de sumoilación. Se ha identificado también la ubiquitinización y acetilación de algunos de los componentes de la maquinaria de SUMO. Diferentes E3 ligasas de SUMO de mamíferos son reconocidas por el sistema de la ubiquitina marcando estas proteínas hacía la degradación vía proteosoma, regulando la disponibilidad de las E3 ligasas de SUMO (Depaux et al., 2007). Por otro lado, cabe destacar la regulación a nivel de sustrato. Uno de los mecanismos más descrito es la fosforilación del sustrato que puede regular positivamente o negativamente la sumoilación de la proteína diana.

OBJETIVOS

El objetivo general de esta tesis es el estudio de las bases moleculares de la función de SUMO en plantas.

Con este fin, se han abordado los siguientes objetivos específicos:

1. Análisis molecular comparativa de los diferentes parálogos de SUMO de *Arabidopsis thaliana*.

2. Caracterización bioquímica y funcional de las isoformas de la enzima activadora de SUMO de *A.thaliana*.

3. Desarrollo de una herramienta molecular con el fin de inhibir la SUMOilación en plantas.
    3.1 Validación *in vitro* e *in vivo*de la estrategia.
    3.2Aplicación de esta herramienta molecular para el análisis del papel de SUMO en respuesta a patógenos necrótrofos.

4. Estudio de las modificaciones post-traduccionales de la enzima activadora de SUMO y sus posibles implicaciones en la regulación de esta enzima.

5. Identificación de putativos sustratos de SUMO durante las primeras etapas de la germinación de la semilla mediante espectroscopia de masas.

CAPÍTULO-I

Distinctive properties of *Arabidopsis* SUMO paralogues support the in vivo predominant role of AtSUMO1/2 isoforms

Distinctive properties of *Arabidopsis* SUMO paralogues support the in vivo predominant role of AtSUMO1/2 isoforms

Laura Castaño-Miquel[1], Josep Seguí[1] and L. Maria Lois[1]

1 Department of Molecular Genetics, Center for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB, Parc de Recerca UAB, Edifici CRAG, Campus UAB, 08193 Bellaterra Barcelona, Spain

Biochem J. 2011 Jun 15;436(3):581-90. doi: 10.1042/BJ20101446.

PMID: 21413927 [PubMed – indexed for MEDLINE]

Resumen. La modificación post-traduccional de las proteínas mediada por SUMO es un mecanismo de regulación crucial en eucariotas. Los mecanismos moleculares del sistema de SUMOilación se encuentran conservados, aunque el número de componentes de la maquinaria de SUMO y su grado de conservación es específico de cada organismo. En el presente trabajo, se muestra que los parálogos de SUMO en *Arabidopsis thaliana*, *At*SUMO1, 2, 3 y 5, han divergido funcionalmente más que sus ortólogos humanos. Se ha analizado el grado de conservación de los parálogos de SUMO, los resultados muestran que las superficies implicadas en la interacción no covalente con la enzima activadora de SUMO (E1), la enzima conjugadora de SUMO (E2) y los motivos SIM no se encuentran conservados entre las diferentes isoformas de SUMO. Los residuos implicados en esta interacción no covalente se encuentran bien conservados en *At*SUMO1/2, mientras que *At*SUMO3 muestra un menor grado de conservación, y *At*SUMO5 presenta una superficie de interacción más divergente. Estas diferencias son funcionalmente relevantes, *At*SUMO3 y 5 muestran deficiencias en la interacción no covalente con la enzima E2 de SUMO, mientras que todos los ortólogos humanos de SUMO son capaces de interaccionar no covalentemente con la enzima E2. Además, se han observado diferencias en la tasa de conjugación de cada uno de los parálogos de SUMO; *At*SUMO3 es conjugado con menos eficiencia que *At*SUMO1/2, y *At*SUMO5 muestra el nivel más bajo de conjugación *in vitro*. Un análisis de mutagénesis dirigida ha revelado que la disminución en la tasa de conjugación y formación del enlace tioéster son consecuencia de los residuos no conservados que participan en la interacción no covalente con la enzima E1. En general, nuestros resultados proporcionan la evidencia de que existe una conjugación preferencial de las isoformas esenciales *At*SUMO1/2 que está modulada por el papel de la enzima activadora en la discriminación de los parálogos de SUMO.

BJ www.biochemj.org

BJ Plant

Biochem. J. (2011) **436**, 581–590 (Printed in Great Britain) doi:10.1042/BJ20101446

online data

Distinctive properties of *Arabidopsis* SUMO paralogues support the *in vivo* predominant role of AtSUMO1/2 isoforms

Laura CASTAÑO-MIQUEL, Josep SEGUÍ and L. Maria LOIS[1]

Department of Molecular Genetics, Center for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB, Parc de Recerca UAB, Edifici CRAG, Campus UAB, 08193 Bellaterra Barcelona, Spain

Protein modification by SUMO (small ubiquitin-related modifier) has emerged as an essential regulatory mechanism in eukaryotes. Even though the molecular mechanisms of SUMO conjugation/deconjugation are conserved, the number of SUMO machinery components and their degree of conservation are specific to each organism. In the present paper, we show data contributing to the notion that the four expressed *Arabidopsis* SUMO paralogues, AtSUMO1, 2, 3 and 5, have functionally diverged to a higher extent than their human orthologues. We have explored the degree of conservation of these paralogues and found that the surfaces involved in E1-activating enzyme recognition, and E2-conjugating enzyme and SIM (SUMO-interacting motif) non-covalent interactions are well conserved in AtSUMO1/2 isoforms, whereas AtSUMO3 shows a lower degree of conservation, and AtSUMO5 is the most divergent isoform. These differences are functionally relevant, since AtSUMO3 and 5 are deficient in establishing E2 non-covalent interactions, which has not been reported for any naturally occurring SUMO orthologue. In addition, AtSUMO3 is less efficiently conjugated than AtSUMO1/2, and AtSUMO5 shows the lowest conjugation level. A mutagenesis analysis revealed that decreases in conjugation rate and thioester-bond formation are the result of the non-conserved residues involved in E1-activating enzyme recognition that are present in AtSUMO3 and 5. The results of the present study support a role for the E1-activating enzyme in SUMO paralogue discrimination, providing a new mechanism to favour conjugation of the essential AtSUMO1/2 paralogues.

Key words: conjugation rate, E1-activating enzyme, E2-conjugating enzyme, non-covalent interaction, poly-small ubiquitin-related modifier (SUMO) chain, small ubiquitin-related modifier (SUMO).

INTRODUCTION

In plants, regulation of protein activity by SUMO (small ubiquitin-related modifier) attachment is a post-translational modification that has been shown to be essential during seed development and to have a major role in abiotic and biotic stress responses [1]. A common property between plants and animals is that the SUMOylation system appears to be a target for pathogenic effectors [2–5], as well as the accumulation of SUMO conjugates in response to heat and oxidative stresses [6,7]. But there are also biological processes specific to plants in which SUMO has a relevant role, such as flowering [8], phosphate starvation [9], drought responses [10] and the abscisic acid signalling pathway [11,12], a hormone that mediates plant responses to environmental stresses and a key regulator of plant growth and development.

SUMO is covalently attached to target proteins by the sequential action of E1-activating, E2-conjugating and E3-ligase enzymes [13]. SUMO activation is mediated by a heterodimeric enzyme consisting of a large subunit, SAE (SUMO-activating enzyme) 2, and a small subunit, SAE1. The SAE2 subunit contains the adenylation, catalytic cysteine, UFD (ubiquitin fold) and C-terminal functional domains [14]. The adenylation domain is responsible for SUMO recognition and adenylation of the C-terminus of SUMO. After adenylation, the catalytic cysteine thiol group attacks the SUMO C-terminal adenylate resulting in the formation of a thioester bond between the E1 and SUMO, in a mechanism that involves a rotation of the cysteine domain [15]. At this stage, SUMO can be transferred to the E2-conjugating enzyme. SUMO-charged E2 is competent to transfer SUMO to the target lysine residue in the substrate, although this reaction is facilitated by E3-ligase enzymes *in vitro* and *in vivo* [16,17].

During conjugation, SUMO molecules establish non-covalent interactions with the E1-activating and the E2-conjugating enzymes. Structural studies have identified eleven residues on the HsSUMO (human SUMO) surface that establish contacts with the E1-activating enzyme and which, presumably, are responsible for conferring modifier specificity [14]. SUMO can be attached to the target protein as a monomer or polymer, and polySUMO chains have been shown to act as signals to promote ubiquitination of the SUMO-modified substrate targeting it for proteasomal degradation [18]. SUMO chain growth is dependent on the presence of a SUMOylation consensus site at the SUMO N-terminal tail, and polymerization is facilitated by non-covalent interactions between SUMO and the E2-conjugating enzyme [19,20]. *In vitro*, the chain length is modulated by the relative abundance of HsSUMO2/3, which can build SUMO chains, compared with HsSUMO1, which does not have a SUMOylation consensus site and that could function as a chain terminator [21].

The consequences of covalent SUMO attachment to target proteins are very variable and include regulation of subcellular localization, protein activity and stability, and protein–protein interactions. At the molecular level, the SUMOylation outcome is achieved through the interaction with specific effectors that contain a SIM (SUMO-interaction motif). Most SIMs consist of a hydrophobic core of three to four aliphatic residues flanked by acidic residues [22,23]. Structural and functional studies determined that a hydrophobic groove surrounded by basic

Abbreviations used: AtCAT3, *Arabidopsis* catalase isoform 3; DTT, dithiothreitol; GST, glutathione transferase; SD, synthetic dextrose; SUMO, small ubiquitin-related modifier; AtSUMO, *Arabidopsis* SUMO; HsSUMO, human SUMO; SAE, SUMO-activating enzyme; AtSAE, *Arabidopsis* SAE; AtSCE1, *Arabidopsis* SUMO-conjugating enzyme 1; SIM, SUMO-interacting motif; ULP, ubiquitin-like protein-specific protease.

[1] To whom correspondence should be addressed (email maria.lois@cid.csic.es).

residues in SUMO is crucial for SIM interaction. Interestingly, the relative position between the hydrophobic groove and the basic residues differ among HsSUMO isoforms and this could confer SUMO–SIM interaction specificity [24,25]. During conjugation, SUMO paralogue selection can be mediated by SIM-dependent recruitment of targets to SUMO thioester-charged E2 and/or SUMO-modified E2 [26,27]. Moreover, SIMs have been identified in E3 ligases and shown to regulate ligase activity and localization [28–30].

In plants, much less is known about the molecular mechanisms that regulate SUMOylation, and the complexity of the SUMOylation components is apparently higher when compared with other organisms. In *Arabidopsis*, expression has been detected for the four SUMO paralogues AtSUMO1, AtSUMO2, AtSUMO3 and AtSUMO5. Among them AtSUMO1 and 2 are the most closely related isoforms. Previous studies have shown that *Arabidopsis* SUMO paralogues do not serve as equivalent substrates of AtSUMO proteases, referred to as ULP (ubiquitin-like protein-specific protease). The four proteases AtULP1a, c, d and AtESD4 (*Arabidopsis* early is short days 4) displayed similar peptidase and isopeptidase activities towards AtSUMO1 and AtSUMO2 isoforms, although none of them showed a significant activity towards AtSUMO5. Only AtULP1a exhibited a poor peptidase activity towards AtSUMO3 and no isopeptidase activity at all [31–33]. In addition, in non-quantitative assays, only AtSUMO1, 2 and 3 were shown to be conjugated to the yeast substrate PCNA (proliferating-cell nuclear antigen), and the capacity to form polymeric chains was displayed exclusively by AtSUMO1 and AtSUMO2 [32]. At first glance, it seemed that the situation of the AtSUMO1 and 2 isoforms resembles that of the human SUMO2 and 3 isoforms according to their capacity to form polymeric chains, as opposite to the HsSUMO1 and the AtSUMO3 isoforms that are conjugated as monomers. But, in contrast, homology studies failed to cluster human and *Arabidopsis* isoforms according to their ability or inability to polymerize [11,32]. In addition, AtSUMO3 and 5 are not capable of complementing the lethal double-mutant *atsumo1atsumo2* plants [34], whereas SUMO1-knockout mice are viable. These results suggest that mammalian SUMO1 and SUMO2/3 have partially redundant functions [35,36], in contrast with AtSUMO paralogues that seem to have developed more divergent functions [34,37].

To gain new insights into the complex *Arabidopsis* SUMOylation system, we have assessed whether AtSUMO1, 2, 3 and 5 have distinct molecular properties that might influence their *in vivo* conjugation and biological function. We have found that *Arabidopsis* SUMO isoforms have heterogeneous properties at different molecular levels. AtSUMO1 and 2 were competent to interact non-covalently with their cognate E2-conjugating enzyme AtSCE1 (*Arabidopsis* SUMO-conjugating enzyme 1), whereas AtSUMO3 and AtSUMO5 did not conserve this property. Mutagenesis analysis revealed that the single residue Asp^{63}, conserved in the AtSUMO1/2 surface but not in AtSUMO3 and AtSUMO5, is essential for non-covalent interactions with E2, and that it is necessary for polySUMO chain formation. It is even more significant the fact that SUMO isoforms differed in their conjugation rate, AtSUMO1/2 being the most efficiently conjugated paralogues, AtSUMO3 was less efficiently conjugated and AtSUMO5 showed the lowest conjugation level. A mutagenesis analysis showed that the lower conjugation rates of AtSUMO3 and AtSUMO5 were related to changes in the SUMO residues involved in the E1 interaction, which also affected thioester-bond formation. These results suggest that the first step in the SUMO-conjugation cascade would have a regulatory role in SUMO paralogue discrimination. Overall, the results of the present study suggest that AtSUMO1/2 might be the most efficiently conjugated SUMO isoforms *in vivo*, and we postulate that this could constitute a molecular mechanism to assure conjugation of the essential AtSUMO1/2 paralogues compared with the non-essential AtSUMO3 and 5.

EXPERIMENTAL

Cloning, expression and protein purification

AtSUMO isoforms, AtSAE2, AtSAE1a/b and AtCAT3 (*Arabidopsis* catalase isoform 3) were amplified by PCR from cDNA obtained from 2-week-old plants grown on MS (Murashige and Skoog) plates under LD (light/dark) at 22 °C (Superscript® III reverse transcriptase from Invitrogen and Pfu DNA polymerase from Stratagene). AtSAE1a/b were cloned into pET15b (Novagen) to encode a native polypeptide, and AtSAE2, AtSCE1 [11], AtSUMO1-(1–93), AtSUMO2-(1–92), AtSUMO3-(1–93) and AtSUMO5-(1–103) were cloned into pET28a (Novagen) to encode an N-terminal His_6-fusion protein. AtCAT3Ct-(419–472) (Ct is C-terminal tail) was cloned into pGEX-6P1 (Amersham) to encode an N-terminal GST (glutathione transferase)-fusion protein. Plasmids were transformed individually, or co-transformed in the case of AtSAE2- and AtSAE1a/b-containing plasmids, into *Escherichia coli* strain BL21 Codon Plus RIL (Stratagene). Cultures (1–4 litres) were incubated at 37 °C until they reached an A_{600} of 0.6–0.8, and protein expression was induced by adding 0.1 mM IPTG (isopropyl β-D-thiogalactopyranoside) for 4 h at 30 °C. Cells were harvested and resuspended in 50 mM Tris/HCl (pH 8.0), 20 % (w/v) sucrose, 350 mM NaCl, 20 mM imidazole, 0.1 % Nonidet P40, 1 mM PMSF, 1 mM 2-mercaptoethanol, 1 μg/ml leupeptin, 1 μg/ml pepstatin and 50 μg/ml DNAse. Protein extracts were prepared by sonication [five pulses of 1 min and 2–3 power setting (Sonifier 250, Branson)] and clarified by centrifugation (18 000 ***g*** for 30 min at 4 °C). Purification via IMAC–Sepharose resin (GE Healthcare) or glutathione–Sepharose (GE Healthcare) was performed according to the manufacturer's instructions. SDS/PAGE analysis of the purified proteins is shown in Supplementary Figure S1 (at http://www.BiochemJ.org/bj/436/bj4360581add.htm).

Yeast two-hybrid experiments

Yeast expression constructs pGBKT7:AtSCE1/AtUBC10 and pGADT7:AtSUMO1-(1–93)/AtUBI have been generated previously [11]. AtSUMO3-(1–93) and AtSUMO5-(1–103) were cloned into pGADT7 AD (Clontech) to encode an N-terminal GAL4-activation domain fusion protein. AtSUMO1 and AtSUMO3 mutant alleles were generated by QuikChange® site-directed mutagenesis (Stratagene). Plasmids, as indicated, were co-transformed into the yeast strain HF7c using the lithium acetate method as described in the Clontech Yeast Protocols Handbook. Transformed yeast culture was plated on to permissive SD (synthetic dextrose) medium complemented with histidine. A single clone per transformation was selected, disaggregated by vortex agitation in SD medium without amino acids, and serial dilutions were performed (1, 1:8, 1:32 and 1:64). Aliquots (5 μl) of each dilution were sowed on non-selective (SD medium complemented with histidine) or selective (SD medium not complemented with histidine) plates. After incubation for 2 days at 30 °C, protein interactions were analysed using histidine auxotrophy as a selective marker.

In vitro pull-down assay

His–AtSCE1 (100 μM) and AtSUMO1/D63N (25 μM) were incubated in 40 μl of binding buffer [20 mM Tris/HCl (pH 8.0), 50 mM NaCl and 20 mM imidazole] for 5 h at 4 °C. Next, 10 μl of Ni^{2+}-IMAC–Sepharose resin was added to the binding mixture, and the mixture was incubated for 30 min at 4 °C. The binding mixture was transferred to micro bio-spin chromatography columns (Bio-Rad) and the resin was washed four times with 20 μl of binding buffer. The proteins bound to the resin were eluted with 20 μl of binding buffer containing 300 mM imidazole. The input (0.8 μl) and eluate (3.5 μl) fractions were separated by SDS/PAGE and either stained with Coomassie Fluor Orange (Molecular Probes, C-33250) or subjected to immunoblot analysis with anti-SUMO1 antibodies, as indicated.

Polyclonal anti-AtSUMO1 antibody production

Polyclonal antibodies were raised against the purified N-terminal His_6-fusion protein of AtSUMO1-(1–93). Purified protein (1 mg) was resolved by SDS/PAGE, the gel was stained with Coomasie Blue and the gel slice containing His–AtSUMO1 was used to immunize rabbits (Cocalico Biological). A 1:1000 dilution of the serum produced was used in the immunoblot analyses.

In vitro SUMO conjugation, polymeric chain formation and E1-thioester assays

In conjugation assays, we used the C-terminal tail of the AtCAT3 (residues 419–472) fused to GST, GST–AtCAT3Ct. Reactions were carried out at the indicated temperatures in 25 μl reaction mixture volumes containing 1 mM ATP, 50 mM NaCl, 20 mM Hepes (pH 7.5), 0.1 % Tween 20, 5 mM $MgCl_2$, 0.1 mM DTT (dithiothreitol), 2 μM SUMO, 0.5 μM AtSAE2/AtSAE1a, 0.5 μM AtSCE1 and 5 μM GST–AtCAT3Ct. After the specified incubation time, reactions were stopped by the addition of protein-loading buffer, boiled for 5 min and 10 μl aliquots were resolved by SDS/PAGE. Polymeric chain formation reactions were performed at 37 °C in the same reaction buffer as SUMO-conjugation assays and in the presence of 100 μM AtSUMO1, 1 μM AtSAE2/AtSAE1a and 10 μM AtSCE1. Reaction products were detected by immunoblot analysis with anti-GST polyclonal antibodies (Sigma, G7781) or with anti-AtSUMO1 polyclonal antibodies, as indicated. E1-thioester assays were performed at 30 °C in 50 μl reaction mixture volumes containing 1 mM ATP, 50 mM NaCl, 20 mM Hepes (pH 7.5), 0.1 % Tween 20, 5 mM $MgCl_2$, 0.1 mM DTT, 10 μM SUMO and 5 μM AtSAE2/AtSAE1a. At the indicated time points, 15 μl aliquots were removed and analysed by SDS/PAGE followed by Coomasie Fluor Orange staining according to manufacturer's protocol (Molecular Probes, C-33250). As a thioester-bond formation control, an aliquot of each reaction was treated with 100 mM DTT before loading on to polyacrylamide gels.

SUMO-conjugation rate quantification

Reaction products were detected using ECL (enhanced chemiluminescence) and Western blot detection reagents (GE Healthcare), and the signal was acquired with the LAS-3000 imaging system and quantified with Multi Gauge V3.0 (Fujifilm). Signals were normalized against known amounts of GST included in each blot. When data are represented by relative units, SUMO-conjugation or E1-thioester rates are referred to the average value calculated using all the rates obtained in each independent experiment.

Bioinformatics

Sequence alignments were performed using the protein multiple alignment software MUSCLE [38] and alignments were edited with GeneDoc software (www.psc.edu/biomed/genedoc). Protein structure models were generated using the SWISS-MODEL workspace [40] on automated mode. AtSUMO1/3/5 and AtSCE1 models were generated using PDB code 2PE6 [2.40 Å (1 Å = 0.1 nm)] or PDB code 2IY1 (2.46 Å) as templates. Models were assembled and images were generated using PyMOL (http://www.pymol.org).

Accession numbers

The assigned accession numbers for the genes studied are as follows: At4g26840 (AtSUMO1), At5g55160 (AtSUMO2), At5g55170 (AtSUMO3), At2g32765 (AtSUMO5), At2g21470 (AtSAE2), At4g24940 (AtSAE1a), At5g50580 (AtSAE1b), At3g57870 (AtSCE1), At4g02890 (AtUBI), At5g53300 (AtUBC10) and At1g20620 (AtCAT3).

RESULTS

Conservation of E1-, E2- and SIM-interacting residues among AtSUMO paralogues

The overall degree of homology among SUMO paralogues ranges from 83 % sequence identity between AtSUMO1 and AtSUMO2, to 42 % and 30 % sequence identity between AtSUMO1 and AtSUMO3 or AtSUMO5 respectively. Interestingly, these differences are also present in the degree of conservation found between residues involved in E1 and E2 non-covalent interactions. AtSUMO1 and AtSUMO2 have identical amino acid residues at the positions involved in non-covalent interactions with the E1-activating and E2-conjugating enzymes. In contrast, AtSUMO3 and AtSUMO5 have more divergent interacting surfaces. AtSUMO3 has 55 % and 75 % degree of conservation of E1- and E2-interacting residues respectively, whereas AtSUMO5 is the most divergent isoform, showing a degree of conservation of 36 % and 56 % for E1- and E2-interacting residues (Figures 1A and 1B, and Supplementary Figure S2 at http://www.BiochemJ.org/bj/436/bj4360581add.htm). Previous structural studies identified 11 residues in the HsSUMO1 surface involved in E1 interactions [14]. Taken as reference positions in AtSUMO1, the four residues, Gln^{25}, Gly^{27}, Gly^{92} and Gly^{93}, are identical in all *Arabidopsis* paralogues. Among the others, two are divergent only in the AtSUMO5 isoform, Arg^{66} and Asp^{85}, and five are not conserved in either AtSUMO3 or AtSUMO5, Asn^{56}, Met^{87}, His^{89}, Gln^{90} and Thr^{91} (Figures 1A and 5A). All of them are identical between AtSUMO1 and 2. With regard to the residues involved in E2 non-covalent interactions, those residues establishing lateral chain contacts with the E2-conjugating enzyme are identical across AtSUMO1/2, HsSUMO2/3 and yeast ySmt3, except for the residue Met^{87} in AtSUMO1, which is highly variable (Figures 1B and 2A). The deduced consensus motif considering the most frequent residues would be Asp^{63}/Glu^{63}, Glu^{79}, Asp^{82} and Asp^{85}/Glu^{85} (the residue numbering is for AtSUMO1). When we analysed AtSUMO3, it is remarkable that the acidic Asp^{63}/Glu^{63} residue in the consensus sequence is substituted with a polar asparagine residue. This change is also present in the most divergent SUMO isoform AtSUMO5 that, in addition, has substitutions in Asp^{77}/Glu^{77} and Asp^{85}/Glu^{85} by histidine and cysteine respectively (Figures 1B and 2A).

Another important interacting surface on SUMO involves the second β-sheet and the downstream α-helix, which form a hydrophobic groove flanked by basic residues that accommodates

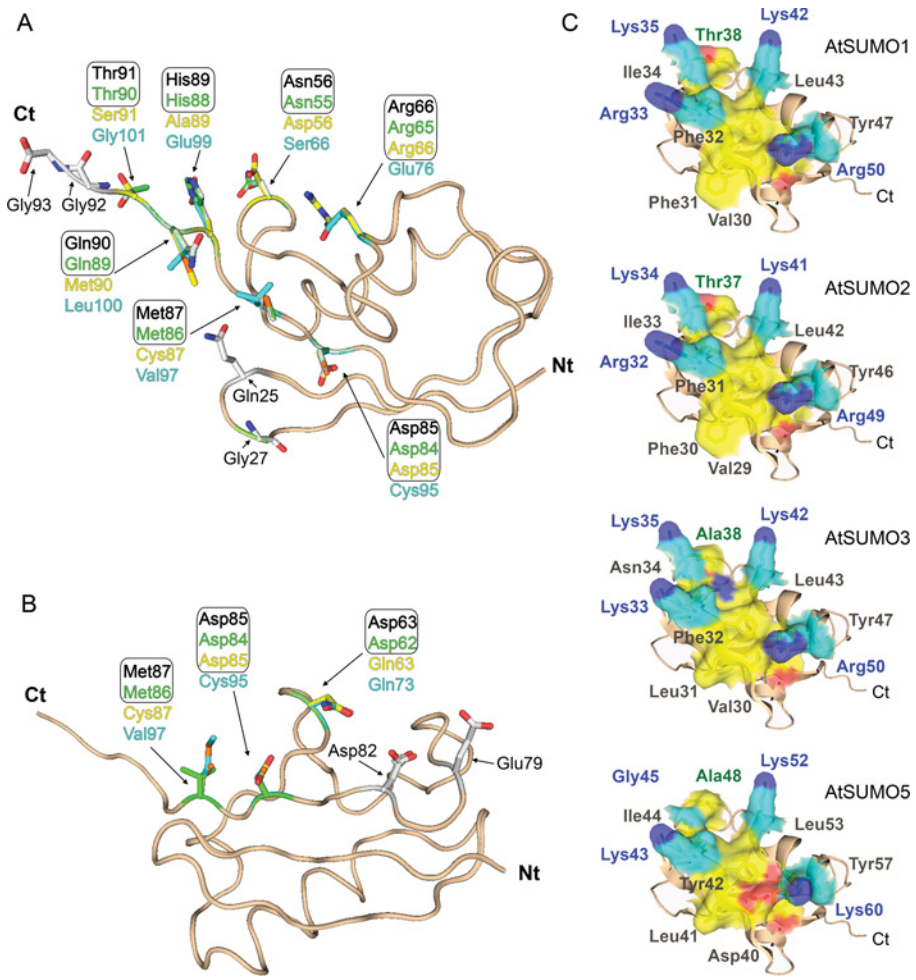


Figure 1 Conservation analysis of residues on AtSUMO paralogue surfaces involved in non-covalent interactions

AtSUMO paralogue structures, aspredicted by the SWISS-MODEL comparative protein modelling server, are shown as a ribbon representations. Residues contained in the N-terminus tail and those absent in the mature forms were not included in the modelling. Residues involved in E1 (**A**) and E2 (**B**) non-covalent interactions are shown in stick representation. Residues corresponding to AtSUMO1, AtSUMO2, AtSUMO3 and AtSUMO5 are shown in black, green, yellow and blue respectively. (**C**) Same structure model as in (**A**) but showing residues involved in SIM interactions in surface representation. Basic residues and the groove-forming residues are coloured cyan and yellow respectively. Side-chain nitrogen and oxygen atoms are coloured blue and red.

SIMs [42]. Structural studies have determined that aliphatic and aromatic residues constitute this hydrophobic groove, HsSUMO1 Ile^{34}, His^{35}, Phe^{36}, Val^{38}, Leu^{47} and Tyr^{51} [23]. Whereas functional studies identified residues required for the role of HsSUMO2 in transcriptional inhibition, which include the hydrophobic Val^{30} and Ile^{34} (equivalent to Ile^{34} and Val^{38} in HsSUMO1), the polar Thr^{38} (Thr^{42} in HsSUMO1), and the four basic residues Lys^{33}, Lys^{35}, Lys^{42} and Arg^{50} (Lys^{37}, Lys^{39}, Lys^{46} and Arg^{54} in HsSUMO1) [24]. Among the basic residues, HsSUMO1 Lys^{39} has been proposed to interact with phosphorylated residues located next to the hydrophobic core, which is the essential component of SIMs [22]. The previous functional amino acids identified in HsSUMO1 and HsSUMO2 are conserved in AtSUMO1/2, suggesting that these paralogues will share the molecular basis for the SIM interaction as their human orthologues. In contrast, major changes are present in AtSUMO3 and AtSUMO5. The hydrophobic Val^{30} and Ile^{34} residues shown to be necessary for transcriptional repression in HsSUMO2 are substituted by an acidic residue in AtSUMO5 (Asp^{40}) and a polar residue (Asn^{34}) in AtSUMO3 respectively. In addition, the polar Thr^{38} in HsSUMO2, which is also necessary for transcriptional repression, is substituted by the hydrophobic alanine residue in both AtSUMO3 and AtSUMO5 (positions 38 and 48 respectively). Finally, the basic Lys^{39} in HsSUMO1 proposed to interact with phosphorylated residues in the target is substituted by an uncharged glycine residue in AtSUMO5 (Figure 1C). Considering that no SIM-containing targets have been identified in *Arabidopsis*, we have focused on the functional analysis of E1 and E2 non-covalent interacting residues in *Arabidopsis* SUMO paralogues, according to their role in conjugation.

E2 non-covalent interactions with SUMO isoforms

In order to assess the effect of changes in residues involved in SUMO–E2 non-covalent interactions, we performed yeast two-hybrid assays. In these experiments, the capacity of the yeast strain HF7c to grow in the absence of histidine was used as a marker for the interaction between proteins. Previous studies have shown that AtSUMO1 and 2 were capable of interacting with AtSCE1 in similar assays [11]. We found that histidine auxotrophy was restored only when AtSCE1a was co-transformed with AtSUMO1, but not with AtSUMO3 or AtSUMO5. In these experiments, AtSUMO1 was used as a positive control and *Arabidopsis* ubiquitin was used as a negative control. When AtSCE1 was replaced by the ubiquitin-conjugating enzyme AtUBC10, we observed that histidine auxotrophy was

conferred only when ubiquitin was co-expressed, consistent with the specificity of the system (Figure 2B). The results of the present study demonstrate that AtSUMO3/5 are not competent to interact with AtSCE1, suggesting the existence of a change in the SUMO-interacting surface that might be common to both paralogues. As described above, the central aspartate residue in AtSUMO1 and 2 (Asp^{63} and Asp^{62}) is replaced by an asparagine residue at the equivalent position in AtSUMO3 and 5 (Asn^{63} and Asn^{73}) (Figures 1B and 2A). To test the role of this divergent residue, we generated the mutants AtSUMO1/D63N and AtSUMO3/D63N and assessed their capacity to interact with AtSCE1. None of these mutant isoforms interacted with AtSCE1 in yeast two-hybrid assays (Figure 2C), indicating that the presence of an aspartate residue at position 63 is essential, but not sufficient, for SUMO–E2 non-covalent interactions. To further evaluate the essential role of Asp^{63} in non-covalent interactions with E2, we performed *in vitro* pull-down assays. In these assays, His–AtSCE1 was incubated in the presence of AtSUMO1 or AtSUMO1 D63N. After binding to a Ni^{2+}-charged resin, His–AtSCE1 was eluted with imidazole and we observed that AtSUMO1 was co-eluted with His–AtSCE1. In contrast, a small amount of AtSUMO1/D63N was present in the elution fraction to the same extent as in the negative control, where His–AtSCE1 was omitted.

Identification and validation of AtCAT3 as a *bona fide* SUMO target *in vitro*

In order to reconstitute a complete *Arabidopsis* SUMOylation system that allowed us the biochemical characterization of the SUMO paralogues, we aimed to identify an endogenous SUMO target. Since SUMOylation is involved in oxidative stress responses, we analysed whether oxidative stress scavengers could be SUMO targets. We found that AtCAT3 contained a SUMOylation consensus site at its C-terminal domain (Supplementary Figure S3A at http://www.BiochemJ.org/bj/436/bj4360581add.htm). Previous studies have determined that SUMO-conjugation sites are located in an extended structure on the surface of the target protein in order to be accessible to the SUMOylation machinery [43]. To determine the position of the putative AtCAT3 SUMOylation site on the quaternary structure, we performed the AtCAT3 structure prediction using the *Exiguobacterium oxidotolerans* catalase structure as a template (PDB code 2J2M). The model generated model indicated that the putative SUMO acceptor lysine residue Lys^{423} in AtCAT3 is fully exposed at the protein surface (Supplementary Figures S3B and S3C). Validation of AtCAT3 as a SUMO substrate was performed by *in vitro* SUMOylation reactions containing the AtCAT3 C-terminal domain, which includes the predicted SUMOylation site (GST–AtCAT3Ct; Supplementary Figure S3D), in the presence of the reconstituted *Arabidopsis* SUMOylation system, AtSAE2, AtSAE1a, AtSCE1 and AtSUMO1. As a result, we detected SUMO conjugation to AtCAT3Ct in an ATP-dependent manner. In addition, the mutant AtCAT3Ct/K423R was unable to accept SUMO (Supplementary Figure S3D). These results validate AtCAT3 as a SUMO target and identify Lys^{423} as the acceptor site for SUMO modification. The advantage of using AtCAT3Ct as a substrate for *in vitro* reactions as opposed to other targets described in the literature is that it does not require the presence of an E3 ligase in order to be modified, which simplifies the biochemical analysis of SUMO conjugation.

AtSUMO1 Asp^{63} is necessary for efficient polySUMO chain formation

Since AtSUMO1/D63N prevented AtSCE1 non-covalent interactions, we tested whether this mutation affected polySUMO

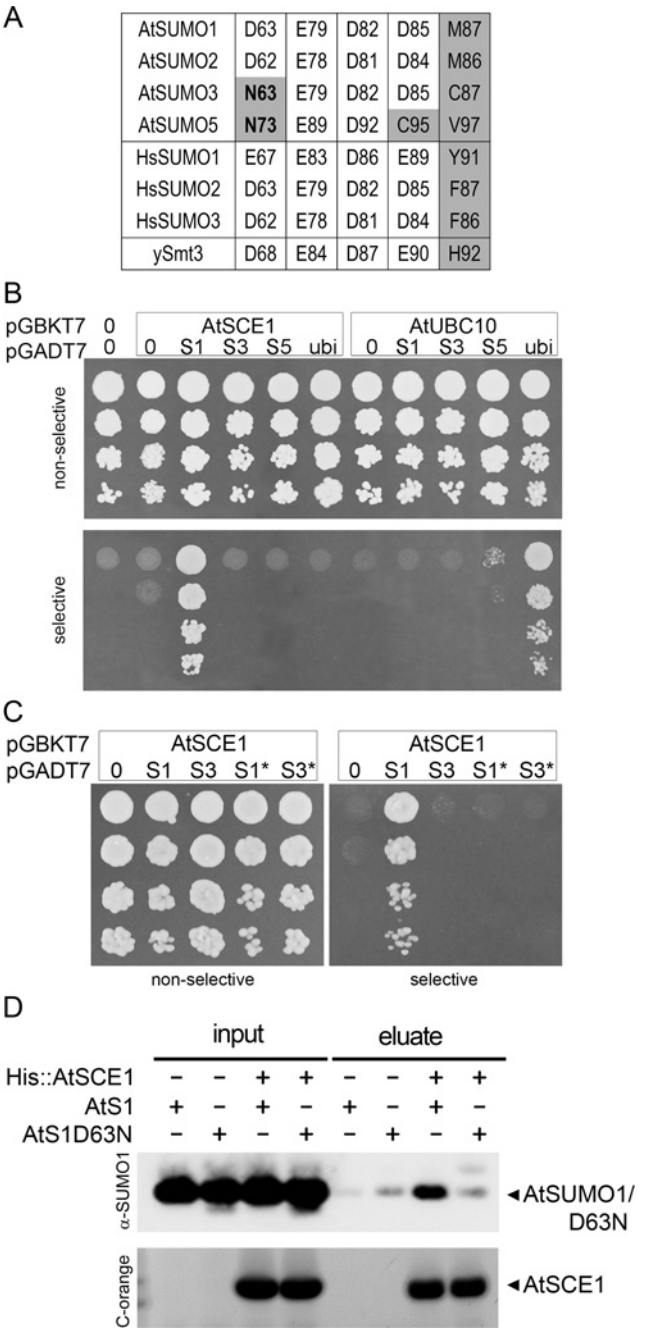


Figure 2 Non-covalent E2-interaction properties of AtSUMO isoforms

(**A**) Sequence alignment of *Arabidopsis* (At), human (Hs) and yeast (y) SUMO residues involved in SUMO–E2 non-covalent interaction through their lateral chain contacts. Non-conserved residues are highlighted in grey, and residues exclusively non-conserved in AtSUMO3/5 are in bold. (**B**) Yeast two-hybrid assay to study interactions between AtSUMO1, 2, 3, 5 or ubiquitin and AtSCE1 or AtUBC10. (**C**) Interaction analysis between the mutant AtSUMO1/D63N (S1*) or AtSUMO3/N63D (S3*) and AtSCE1 as in (**B**). Native SUMO isoforms were also included as a control. (**D**) Poly-histidine pull-down assay of AtSUMO1 or the AtSUMO1/D63N mutant variant using His–AtSCE1 as a bait. Incubations in the absence of the bait or the prey were used as negative controls. Aliquots of input and eluate fractions were resolved by SDS/PAGE. AtSUMO1/D63N and AtSCE1 were analysed by immunoblotting or Coomassie Fluor Orange staining (C-orange) respectively.

chain formation. *In vitro* polySUMO chain-formation assays were performed in the presence of AtSAE2/AtSAE1a, AtSCE1, and the native or mutated SUMO form. Under these conditions, the native AtSUMO1 isoform efficiently built polymeric chains and it was also conjugated to AtSCE1, on which polySUMO chains

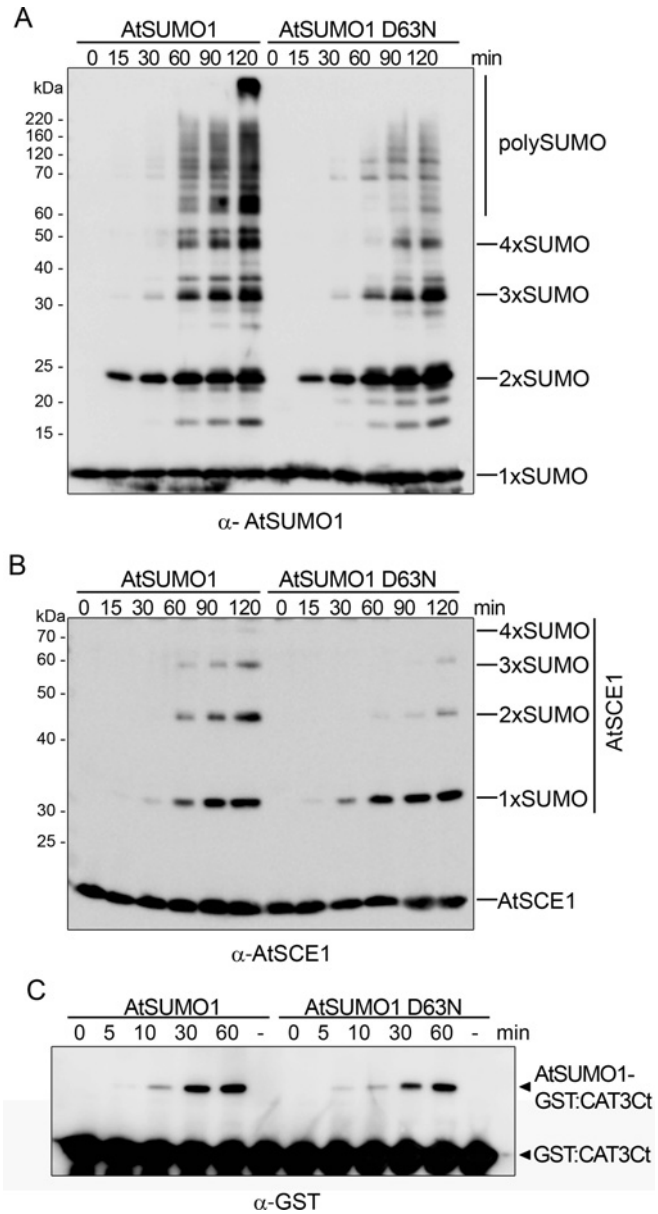


Figure 3 PolySUMO chain formation is dependent on Asp^{63} of AtSUMO1

In vitro SUMOylation assays were performed in the presence of AtSAE2/AtSAE1a, AtSUMO1 or AtSUMO1/D63N and AtSCE1. Reaction mixtures were incubated at 37 °C and stopped at the specified time points. Reaction products were resolved by SDS/PAGE and examined by immunoblot analysis with anti-AtSUMO1 (**A**) or anti-AtSCE1 (**B**) antibodies. (**C**) *In vitro* SUMOylation assays were performed in the presence of AtCAT3Ct as a substrate. Reaction mixtures were incubated at 37 °C, aliquots were removed at the specified time points and reaction products were analysed by immunoblot analysis with anti-GST antibodies. The molecular mass in kDa is indicated on the left-hand side.

were also formed. When Asp^{63} was replaced by an asparagine residue, a reduction in polySUMO chain formation was observed. This defect was more evident from the second conjugation cycle, independently of whether polySUMO chains in the strict sense or polySUMO chains built on AtSCE1 were analysed (Figures 3A and 3B). These results indicate that the mechanism to build polySUMO chains is conserved in *Arabidopsis* and that, presumably, the naturally occurring E2-non-interacting SUMO isoforms AtSUMO3 and AtSUMO5 will not interfere with polyAtSUMO1/2 chain formation *in vivo*. Consistent with a main role for SUMO–E2 non-covalent interactions in polySUMO chain formation, the mutant AtSUMO1/D63N was conjugated to the target AtCAT3Ct with the same efficiency as the native AtSUMO1 (Figure 3C).

SUMO isoforms display a distinct conjugation rate *in vitro*

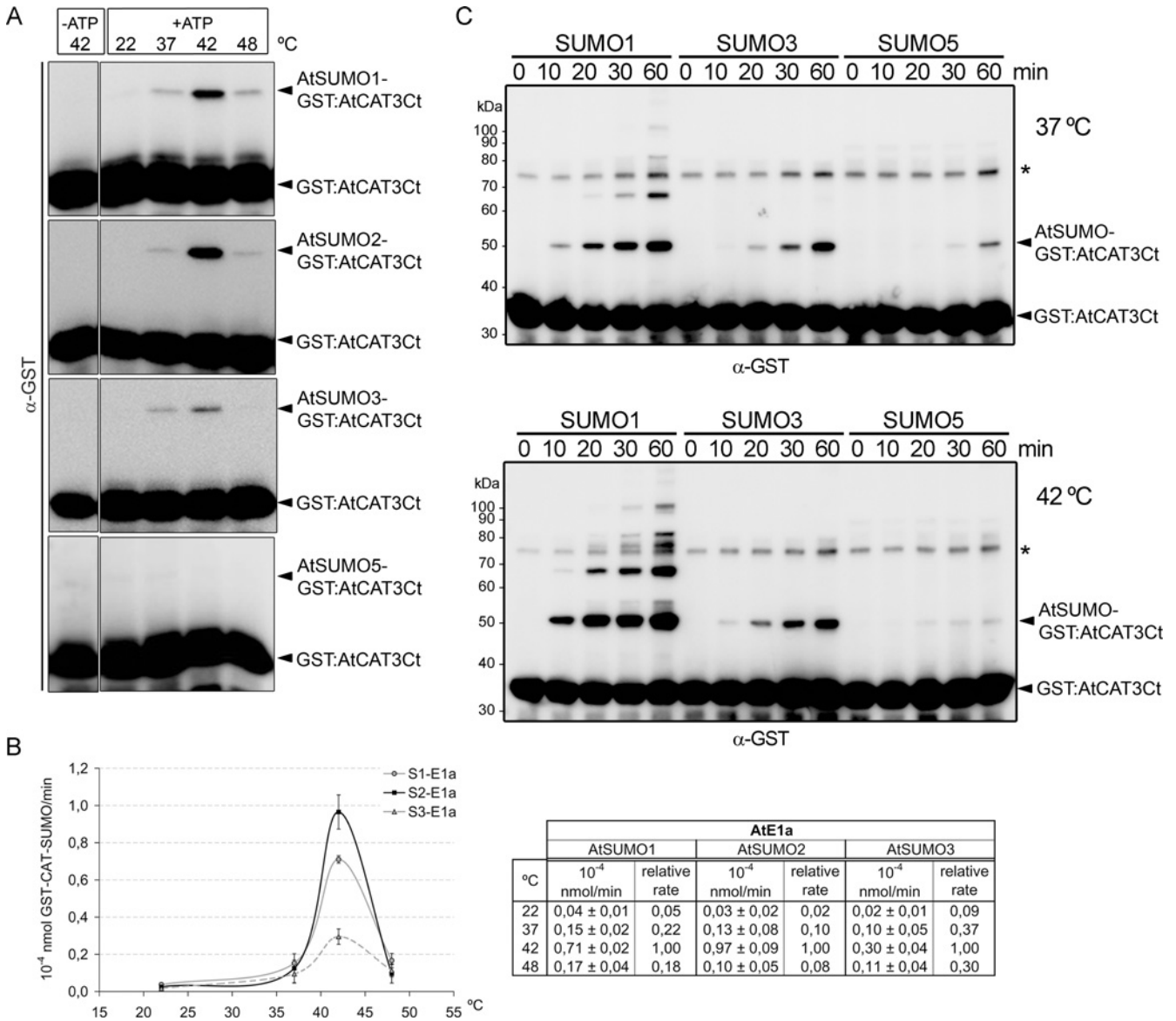
Since AtSUMO isoforms show differences in their capacity to establish non-covalent interactions with their cognate E2-conjugating enzyme, we explored the possibility that they could also differ in their conjugation rates. First, we chose a short incubation time, 10 min, in order to compare the first conjugation cycle. In this way, we avoided conjugation rate underestimation of the isoforms competent for polySUMO chain formation. We were also interested in analysing the effect of the incubation temperature on conjugation rate, which could be biologically relevant since SUMO conjugates accumulate massively upon heat stress. Reaction mixtures were incubated at 22, 37, 42 and 48 °C, and we observed that the highest reaction rate was achieved at 42 °C (Figure 4A). Under these experimental conditions, we did not observe conjugation of AtSUMO5. All other isoforms were conjugated to AtCAT3 at different rates. In general, AtSUMO1 and 2 were better conjugated than AtSUMO3, the highest difference being observed at 42 °C. At this temperature, the AtSUMO1 and AtSUMO2 conjugation rate is 2.4- and 3.2-fold higher than AtSUMO3 respectively (Figure 4B). Next, we performed a time-course analysis that allowed us to detect AtSUMO5 conjugation after 60 min incubation at 37 °C. At 42 °C, AtSUMO5 conjugation was very weak, and differences in the conjugation level between AtSUMO1 and AtSUMO5, as well as between AtSUMO1 and AtSUMO3, were more pronounced than at 37 °C (Figure 4C).

Non-conserved residues involved in the E1 interaction are responsible for differences in the SUMO paralogue conjugation rate

As has been described above, SUMO residues involved in E1 recognition are not evenly conserved between AtSUMO isoforms. In order to evaluate the effect of these changes in conjugation rate, we focused on those positions that were not conserved only in the less conjugated isoform AtSUMO5, which corresponds to AtSUMO1 Arg^{66} and Asp^{85}. We also focused on AtSUMO1 His^{89} which is not conserved either in AtSUMO3 or AtSUMO5, and whose equivalent position in Nedd8 has previously been reported to confer modifier specificity [44]. We mutated residues Arg^{66}, Asp^{85} and His^{89} of AtSUMO1 to those present in AtSUMO3 and AtSUMO5 at the equivalent positions, and analysed their conjugation rates (Figures 5B and 5C). At 42 °C, the most dramatic effect was observed in D85C followed by H89A and R66E mutants, which showed 24 %, 40 % and 50 % of the native AtSUMO1 rate respectively. The H89E substitution had a smaller effect and showed 90 % of the native AtSUMO1 activity. Similar to what we observed during the conjugation analysis of SUMO isoforms, at 37 °C differences in conjugation levels were smaller, although their behaviour was similar, and the D85C mutation was the most affected. These mutants, with the exception of the H89E substitution, also showed SUMO–E1 thioester-bond formation defects, and the D85C mutation was the most affected. These results suggest that the conjugation level reduction of these mutants was the result of E1-interaction defects.

DISCUSSION

The molecular consequences of protein modification by SUMO will be dependent on the specific SUMO paralogue that is attached to the target, and major effort has been put into elucidating the



	AtE1a					
	AtSUMO1		AtSUMO2		AtSUMO3	
°C	10^{-4} nmol/min	relative rate	10^{-4} nmol/min	relative rate	10^{-4} nmol/min	relative rate
22	0,04 ± 0,01	0,05	0,03 ± 0,02	0,02	0,02 ± 0,01	0,09
37	0,15 ± 0,02	0,22	0,13 ± 0,08	0,10	0,10 ± 0,05	0,37
42	0,71 ± 0,02	1,00	0,97 ± 0,09	1,00	0,30 ± 0,04	1,00
48	0,17 ± 0,04	0,18	0,10 ± 0,05	0,08	0,11 ± 0,04	0,30

Figure 4 ***In vitro*** **conjugation properties of AtSUMO isoforms**

(**A**) *In vitro* SUMOylation assays were performed in the presence of AtE1a, AtSUMO1, AtSUMO2, AtSUMO3 or AtSUMO5, AtSCE1a and GST–AtCAT3Ct as a substrate. Reaction mixtures were incubated at the indicated temperature and stopped after incubation for 10 min. Reaction products were resolved by SDS/PAGE and examined by immunoblot analysis with anti-GST antibodies. (**B**) Reactions were performed in triplicate and the GST–AtCAT3Ct SUMOylation rate was quantified. Values are means ± S.D. The Table containing the data plotted into the graph is shown to the right-hand side. (**C**) A time course of SUMO conjugation *in vitro* reactions were performed at 37 °C (top panel) and 42 °C (bottom panel). Reaction mixtures were set up as described in (**A**). The asterisk indicates a contaminating protein. The molecular mass in kDa is indicated on the left-hand side.

mechanisms involved in SUMO paralogue specificity. Among these mechanisms, ULP-mediated deconjugation and SIM-mediated conjugation have been proposed to facilitate SUMO paralogue selection [45]. In the present study we have described results supporting a role of the E1-activating enzyme in SUMO paralogue discrimination in *Arabidopsis*. The analysis of the molecular properties of the four AtSUMO isoforms expressed indicates that they have diverged to a higher degree than their human orthologues, and that the essential AtSUMO1 and AtSUMO2 are the most functionally conserved isoforms.

Non-covalent interactions between SUMO and its cognate E2-conjugating enzyme is an intrinsic property of the system. Initial studies showed that this property is conserved in AtSUMO1 and 2 [11] but, surprisingly, our results showed that AtSUMO3 and 5 do not retain the capacity to interact with their cognate E2, AtSCE1. This situation is unique to the *Arabidopsis* system, since all Hs-SUMO isoforms and yeast Smt3 interact efficiently with their cognate E2-conjugating enzyme. We have also identified AtSUMO1 Asp^{63} as an essential residue for establishing E2 non-covalent interactions. But the fact that AtSUMO3 Asn^{63} substitution with aspartate did not confer competence for this interaction suggests that the inability of AtSUMO3/5 to interact with AtSCE1 could be the result of two types of amino acid changes: the loss of essential residues and the appearance of residues that would be detrimental for this interaction. SUMO–E2 non-covalent interactions have been proposed to be involved in SUMO chain formation of human SUMO2 and 3 isoforms, which also contain a SUMO attachment site in their N-terminal tail [19,20]. In this model, HsSUMO1

A

AtSUMO1	Q25	G27	N56	R66	D85	M87	H89	Q90	T91	G92	G93
AtSUMO2	Q24	G25	N55	R65	D84	M86	H88	Q89	T90	G91	G92
AtSUMO3	Q25	G26	D56	R66	D85	C87	A89	**M90**	S91	G92	G93
AtSUMO5	Q35	G37	S66	E76	**C95**	V97	E99	**L100**	G101	G102	G103
HsSUMO1	Q29	S31	N60	R70	E89	Y91	E93	Q94	T95	G96	G97
HsSUMO2	Q25	G27	R56	P66	D85	F87	Q89	Q90	T91	G92	G93
HsSUMO3	Q24	G26	R55	P65	D84	F86	Q88	Q89	T90	G91	G92
ySmt3		S32	D61	R71	E90	H92	E94	Q95	I96	G97	G98

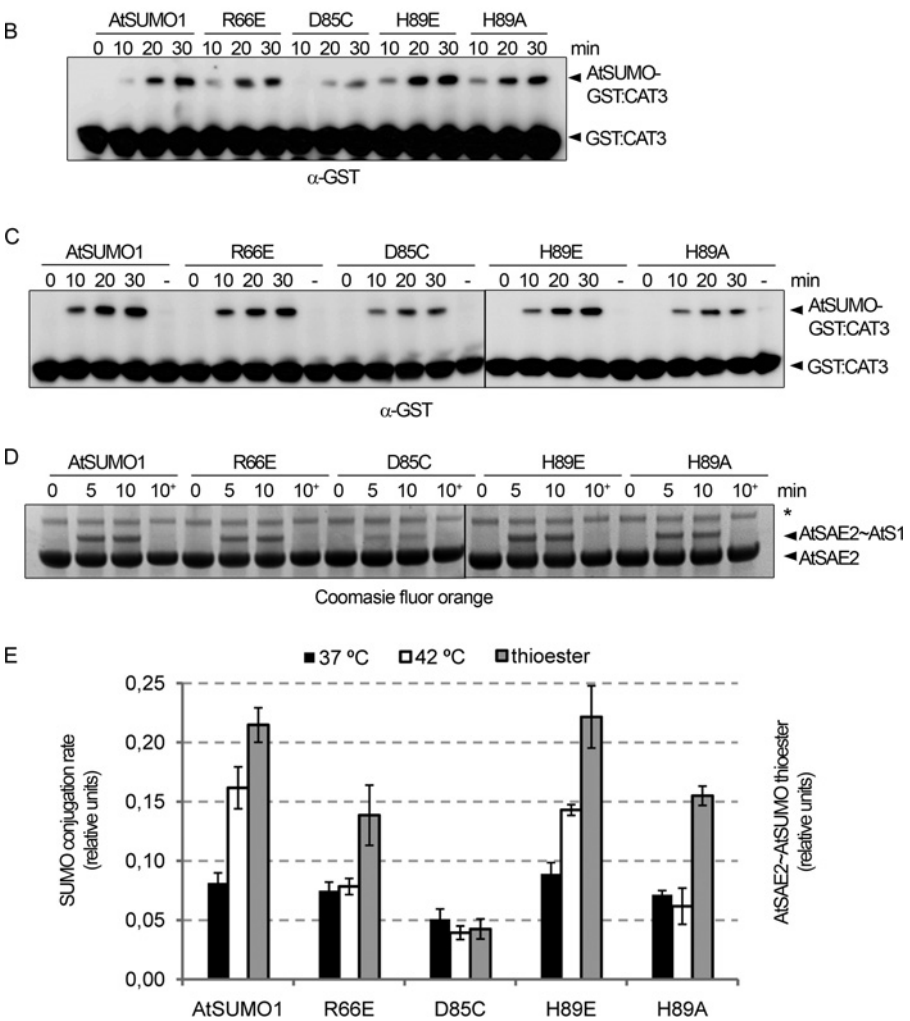


Figure 5 Mutations in the AtSUMO1 residues involved in E1 recognition resulted in SUMO-conjugation defects

(**A**) Sequence alignment of *Arabidopsis* (At), human (Hs) and yeast (y) SUMO residues involved in the SUMO–E1 non-covalent interaction. Non-conserved residues are highlighted in grey and residues exclusively non-conserved in AtSUMO3/5 are in bold. (**B**) *In vitro* SUMOylation assays were performed in the presence of AtE1a, AtSUMO1 or the indicated mutant, AtSCE1a and GST–AtCAT3Ct as a substrate. Reaction mixtures were incubated at 37 °C and stopped at the indicated time points. Reaction products were resolved by SDS/PAGE and examined by immunoblot analysis with anti-GST antibodies. (**C**) The same as in (**B**), but reactions were incubated at 42 °C. (**D**) Time course for AtSAE2–thioester formation using *Arabidopsis* E1 (AtSAE2–AtSAE1a) and AtSUMO1 and AtSUMO1 mutant variants as substrates. Reaction mixtures were incubated at 30 °C and aliquots were removed at the specified time points. In each case, a 10 min aliquot was treated with DTT as a control for thioester-bond formation and this is indicated as 10^{+}. Reaction products were separated by SDS/PAGE and stained with Coomassie Fluor Orange. The asterisk indicates a contaminating protein. (**E**) Reactions were performed in triplicate and GST–AtCAT3Ct SUMOylation and AtSAE2–thioester formation rates were quantified. Average values of the relative activity among AtSUMO1 variants and the corresponding S.E.M. are plotted.

would function as a chain terminator, since it does not comprise an acceptor lysine residue for another SUMO molecule, whereas it retains the capacity to interact non-covalently with E2 enzyme. In *Arabidopsis*, a SUMOylation consensus site is only present in those paralogues capable of interacting non-covalently with AtSCE1, which are AtSUMO1/2, correlating with their ability to form polymeric chains. On the other hand, considering their inability to interact with AtSCE1, it is not clear whether AtSUMO3/5 would have a role as polyAtSUMO1/2 chain terminators, as has been proposed for the HsSUMO1, suggesting that other molecular mechanisms might regulate polySUMO chain length in *Arabidopsis*. Supporting this, recent proteomic studies have failed to identify AtSUMO3/5 peptides in purified AtSUMO1 conjugates [46].

Another remaining question to be addressed was whether AtSUMO paralogues also differed in their conjugation levels. In addition, since SUMO conjugates accumulate dramatically upon heat stress, we were interested in studying the effect of

the temperature in conjugation reactions. In previously reported assays, experiments were designed in such a manner that non-quantitative results were obtained (in most cases the incubation time ranged from a few hours to overnight) [32,47]. The results of the present study, using a quantitative SUMOylation assay, showed that AtSUMO1 and 2 isoforms were more efficiently conjugated in comparison with AtSUMO3, whereas AtSUMO5 showed the lowest conjugation level. For the SUMO isoforms tested, with the exception of AtSUMO5, the conjugation rate increased with temperature, and it was striking to observe that the highest activity occurred at 42 °C. Even though this temperature is higher than standard environmental conditions, it highlights the robustness of the SUMO-conjugation system. At the same time it suggests that the massive and rapid SUMO conjugation observed in plants upon heat-shock treatments could be mediated, at least in part, by the increasing activity of the conjugation system with temperature. In addition, this effect is more pronounced in the case of AtSUMO1 and AtSUMO2 conjugation, which are the isoforms that are greatly conjugated under heat stress.

Differences in conjugation efficiency among SUMO paralogues are also specific to the *Arabidopsis* SUMOylation system, since all three human SUMO isoforms have been shown to form E1- and E2-thioester bonds and be conjugated to the substrate RanGAP with the same efficiency [14], suggesting that, in this case, SUMO paralogue selection will be dependent on availability or SIM-mediated interactions. The conservation analysis of the 11 SUMO residues involved in the E1 adenylation domain interaction [14] showed that five and seven residues are not conserved in AtSUMO3 and AtSUMO5 respectively, suggesting that adenylation could be deficient for these paralogues. A mutagenesis analysis revealed that some of these substitutions were responsible for a lower conjugation rate, although the effect was dependent on the nature of the substitution. The most deleterious was the D85C substitution present in AtSUMO5, supporting the previous result indicating that AtSUMO5 was the less-conjugated isoform. Interestingly, this position in HsSUMO1, Glu^{89}, was previously shown to be crucial for E1-thioester formation [19], supporting our results pointing to a critical role of this residue in SUMO–E1 recognition. As to AtSUMO3, the only substitution tested occurs at a position previously proposed to contribute to modifier discrimination [44] and, when introduced in AtSUMO1, it reduced the conjugation rate to a value equivalent to that of AtSUMO3, suggesting that this substitution has a major contribution in AtSUMO3 conjugation-rate reduction. Conjugation defects in these mutants were more pronounced when reaction mixtures were incubated at higher temperatures, similar to what we observed when comparing SUMO paralogues. Moreover, the reduction in conjugation correlated with a reduction in E1-thioester formation, supporting a function of the residues tested in SUMO recognition by the E1. To our knowledge, this natural occurring SUMO paralogue discrimination by the E1-activating enzyme has not been reported previously in any other system.

Among the different studies that have addressed characterization of the *Arabidopsis* SUMOylation system, *in vitro* AtSUMO5 conjugation has only been detected to a mammalian substrate, RanGAP, in the presence of the mammalian conjugation system [31]. It is well established that SUMO conjugation is highly regulated by specific protein–protein interactions, and using heterologous systems to test SUMO conjugation might give results that could not be relevant *in vivo*. This could explain why AtSUMO5 conjugation was facilitated by the SUMO mammalian system, but when we have used the *Arabidopsis* system, which presumably is more selective, AtSUMO5 was very poorly conjugated. Similarly, a previous report showed that HsSUMO1 was able to interact with the *Arabidopsis* ubiquitin-conjugating enzyme AtUBC10 in yeast two-hybrid assays. In contrast, in the same assays, neither AtSUMO1 nor AtSUMO2 interacted with AtUBC10, suggesting that selective interactions within the SUMO pathway are more permissive when heterologous SUMOylation components are used [11], and highlighting the importance of using homologous systems in biochemical studies.

In vivo, different evidence points to a predominant role for AtSUMO1/2 paralogues. Early studies have shown that endogenous AtSUMO1 and 2 [7,11] and AtSUMO3 [7] were conjugated *in planta*, although AtSUMO3 conjugate levels were lower than for AtSUMO1/2. Instead, endogenous AtSUMO5 conjugation has not been observed. In addition, according to the Genevestigator database, *atsumo3* and *atsumo5* mRNA levels are on average 10-fold lower than *atsumo1/2*. Furthermore, none of the SUMO-specific proteases identified are competent to process AtSUMO5, and only AtULP1a displays an inefficient peptidase activity towards AtSUMO3, suggesting that, even if expressed at lower levels, it would have to be determined which fraction is present in its mature/conjugable form [31,32]. Finally, the fact that double-mutant *atsumo1atsumo2* plants are not viable [34] suggests a biological specialization and, regardless of AtSUMO3 and 5 *in vivo* function, it seems clear that they cannot compensate for the AtSUMO1/2 loss. A recent study suggests that AtSUMO3 function might differ from AtSUMO1 and AtSUMO2 in flowering and salicylic-acid-dependent responses, although homozygous *atsumo3* mutant plants showed normal plant development [37]. Interestingly, a more recent report described the identification of proteins that are specifically conjugated by AtSUMO3 and not AtSUMO1, although the molecular mechanism of this specificity remains elusive [48]. Surprisingly, almost half of the substrates analysed were exclusively modified by AtSUMO3, in contrast with the low levels of AtSUMO3 conjugates previously detected in crude plant extracts.

Overall, different molecular mechanisms seem to converge in order to assure a proper conjugation of the essential *Arabidopsis* SUMO1/2 isoforms compared with the non-essential AtSUMO3/5. These mechanisms comprise regulation of expression levels, maturation and release from targets, and AtSUMO1/2 are the highest expressed isoforms and the most efficient substrates of the characterized endogenous proteases [31,32], suggesting that most of the endogenous pool of mature SUMO will comprise these two isoforms. In the present study, we provide evidence for the existence of a preferential conjugation of AtSUMO1/2 compared with AtSUMO3/5, which is determined by a role of the E1-activating enzyme in SUMO paralogue discrimination.

AUTHOR CONTRIBUTION

Laura Castaño-Miquel performed mutagenesis experiments, characterization of mutant variants, and time-course experiments; Josep Seguí performed the initial characterization of SUMO isoforms; and Maria Lois designed the experiments and wrote the manuscript.

ACKNOWLEDGEMENTS

We thank C.D. Lima for technical advice and M. Rodríguez-Concepción for critical reading of the manuscript before submission.

FUNDING

The work was supported by the European Research Council [grant number ERC-2007-StG-205927], and the Spanish Ministry of Education and Science [grant numbers BIO2008-01495, CONSOLIDER CSD 2007-00036]. L.C.M. is supported by a predoctoral contract through the CRAG. J.S. is supported by a predoctoral fellowship from the Ministry of

Education and Science [grant number BES-2005-6843]. We also thank the 'Departament d'Innovació, Universitats i Empresa' from the Generalitat de Catalunya (Xarxa de Referència en Biotecnologia and 2009SGR 09626) for support.

REFERENCES

1 Lois, L. M. (2010) Diversity of the SUMOylation machinery in plants. Biochem. Soc. Trans. **38**, 60–64
2 Hanania, U., Furman-Matarasso, N., Ron, M. and Avni, A. (1999) Isolation of a novel SUMO protein from tomato that suppresses EIX-induced cell death. Plant J. **19**, 533–541
3 Hotson, A., Chosed, R., Shu, H., Orth, K. and Mudgett, M. B. (2003) *Xanthomonas* type III effector XopD targets SUMO-conjugated proteins in planta. Mol. Microbiol. **50**, 377–389
4 Roden, J., Eardley, L., Hotson, A., Cao, Y. and Mudgett, M. B. (2004) Characterization of the *Xanthomonas* AvrXv4 effector, a SUMO protease translocated into plant cells. Mol. Plant Microbe Interact. **17**, 633–643
5 Castillo, A. G., Kong, L. J., Hanley-Bowdoin, L. and Bejarano, E. R. (2004) Interaction between a geminivirus replication protein and the plant sumoylation system. J. Virol. **78**, 2758–2769
6 Saitoh, H. and Hinchey, J. (2000) Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. J. Biol. Chem. **275**, 6252–6258
7 Kurepa, J., Walker, J. M., Smalle, J., Gosink, M. M., Davis, S. J., Durham, T. L., Sung, D. Y. and Vierstra, R. D. (2003) The small ubiquitin-like modifier (SUMO) protein modification system in *Arabidopsis*. Accumulation of SUMO1 and -2 conjugates is increased by stress. J. Biol. Chem. **278**, 6862–6872
8 Murtas, G., Reeves, P. H., Fu, Y. F., Bancroft, I., Dean, C. and 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
9 Miura, K., Rus, A., Sharkhuu, A., Yokoi, S., Karthikeyan, A. S., Raghothama, K. G., Baek, D., Koo, Y. D., Jin, J. B., Bressan, R. A. et al. (2005) The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. Proc. Natl. Acad. Sci. U.S.A. **102**, 7760–7765
10 Catala, R., Ouyang, J., Abreu, I. A., Hu, Y., Seo, H., Zhang, X. and Chua, N. H. (2007) The *Arabidopsis* E3 SUMO ligase SIZ1 regulates plant growth and drought responses. Plant Cell **19**, 2952–2966
11 Lois, L. M., Lima, C. D. and Chua, N. H. (2003) Small ubiquitin-like modifier modulates abscisic acid signaling in *Arabidopsis*. Plant Cell **15**, 1347–1359
12 Miura, K., Lee, J., Jin, J. B., Yoo, C. Y., Miura, T. and Hasegawa, P. M. (2009) Sumoylation of ABI5 by the *Arabidopsis* SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. Proc. Natl. Acad. Sci. U.S.A. **106**, 5418–5423
13 Ulrich, H. D. (2009) The SUMO system: an overview. Methods Mol. Biol. **497**, 3–16
14 Lois, L. M. and Lima, C. D. (2005) Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. EMBO J. **24**, 439–451
15 Olsen, S. K., Capili, A. D., Lu, X., Tan, D. S. and Lima, C. D. (2010) Active site remodelling accompanies thioester bond formation in the SUMO E1. Nature **463**, 906–912
16 Johnson, E. S. (2004) Protein modification by SUMO. Annu. Rev. Biochem. **73**, 355–382
17 Melchior, F., Schergaut, M. and Pichler, A. (2003) SUMO: ligases, isopeptidases and nuclear pores. Trends Biochem. Sci. **28**, 612–618
18 Geoffroy, M. C. and Hay, R. T. (2009) An additional role for SUMO in ubiquitin-mediated proteolysis. Nat. Rev. Mol. Cell Biol. **10**, 564–568
19 Knipscheer, P., van Dijk, W. J., Olsen, J. V., Mann, M. and Sixma, T. K. (2007) Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. EMBO J. **26**, 2797–2807
20 Capili, A. D. and Lima, C. D. (2007) Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. J. Mol. Biol. **369**, 608–618
21 Matic, I., van Hagen, M., Schimmel, J., Macek, B., Ogg, S. C., Tatham, M. H., Hay, R. T., Lamond, A. I., Mann, M. and Vertegaal, A. C. (2008) *In vivo* identification of human small ubiquitin-like modifier polymerization sites by high accuracy mass spectrometry and an *in vitro* to *in vivo* strategy. Mol. Cell Proteomics **7**, 132–144
22 Hecker, C. M., Rabiller, M., Haglund, K., Bayer, P. and Dikic, I. (2006) Specification of SUMO1- and SUMO2-interacting motifs. J. Biol. Chem. **281**, 16117–16127
23 Song, J., Durrin, L. K., Wilkinson, T. A., Krontiris, T. G. and Chen, Y. (2004) Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. Proc. Natl. Acad. Sci. U.S.A. **101**, 14373–14378
24 Chupreta, S., Holmstrom, S., Subramanian, L. and Iniguez-Lluhi, J. A. (2005) A small conserved surface in SUMO is the critical structural determinant of its transcriptional inhibitory properties. Mol. Cell. Biol. **25**, 4272–4282
25 Song, J., Zhang, Z., Hu, W. and Chen, Y. (2005) Small ubiquitin-like modifier (SUMO) recognition of a SUMO binding motif: a reversal of the bound orientation. J. Biol. Chem. **280**, 40122–40129
26 Meulmeester, E., Kunze, M., Hsiao, H. H., Urlaub, H. and Melchior, F. (2008) Mechanism and consequences for paralog-specific sumoylation of ubiquitin-specific protease 25. Mol. Cell **30**, 610–619
27 Knipscheer, P., Flotho, A., Klug, H., Olsen, J. V., van Dijk, W. J., Fish, A., Johnson, E. S., Mann, M., Sixma, T. K. and Pichler, A. (2008) Ubc9 sumoylation regulates SUMO target discrimination. Mol. Cell **31**, 371–382
28 Merrill, J. C., Melhuish, T. A., Kagey, M. H., Yang, S. H., Sharrocks, A. D. and Wotton, D. (2010) A role for non-covalent SUMO interaction motifs in Pc2/CBX4 E3 activity. PLoS ONE **5**, e8794
29 Cho, G., Lim, Y. and Golden, J. A. (2009) SUMO interaction motifs in Sizn1 are required for promyelocytic leukemia protein nuclear body localization and for transcriptional activation. J. Biol. Chem. **284**, 19592–19600
30 Yang, S. H. and Sharrocks, A. D. (2010) The SUMO E3 ligase activity of Pc2 is coordinated through a SUMO interaction motif. Mol. Cell. Biol. **30**, 2193–2205
31 Chosed, R., Mukherjee, S., Lois, L. M. and Orth, K. (2006) Evolution of a signalling system that incorporates both redundancy and diversity: *Arabidopsis* SUMOylation. Biochem. J. **398**, 521–529
32 Colby, T., Matthai, A., Boeckelmann, A. and Stuible, H. P. (2006) SUMO-conjugating and SUMO-deconjugating enzymes from *Arabidopsis*. Plant Physiol. **142**, 318–332
33 Chosed, R., Tomchick, D. R., Brautigam, C. A., Mukherjee, S., Negi, V. S., Machius, M. and Orth, K. (2007) Structural analysis of *Xanthomonas* XopD provides insights into substrate specificity of ubiquitin-like protein proteases. J. Biol. Chem. **282**, 6773–6782
34 Saracco, S. A., Miller, M. J., Kurepa, J. and Vierstra, R. D. (2007) Genetic analysis of SUMOylation in *Arabidopsis*: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. Plant Physiol. **145**, 119–134
35 Evdokimov, E., Sharma, P., Lockett, S. J., Lualdi, M. and Kuehn, M. R. (2008) Loss of SUMO1 in mice affects RanGAP1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3. J. Cell Sci. **121**, 4106–4113
36 Zhang, F. P., Mikkonen, L., Toppari, J., Palvimo, J. J., Thesleff, I. and Janne, O. A. (2008) SUMO-1 function is dispensable in normal mouse development. Mol. Cell. Biol. **28**, 5381–5390
37 Van Den Burg, H. A., Kini, R. K., Schuurink, R. C. and Takken, F. L. (2010) *Arabidopsis* small ubiquitin-like modifier paralogs have distinct functions in development and defense. Plant Cell **22**, 1998–2016
38 Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**, 1792–1797
39 Reference deleted
40 Arnold, K., Bordoli, L., Kopp, J. and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics **22**, 195–201
41 Reference deleted
42 Kerscher, O. (2007) SUMO junction-what's your function? New insights through SUMO-interacting motifs. EMBO Rep. **8**, 550–555
43 Bernier-Villamor, V., Sampson, D. A., Matunis, M. J. and 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
44 Walden, H., Podgorski, M. S., Huang, D. T., Miller, D. W., Howard, R. J., Minor, Jr, D. L., Holton, J. M. and Schulman, B. A. (2003) The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1. Mol. Cell **12**, 1427–1437
45 Wilkinson, K. A. and Henley, J. M. (2010) Mechanisms, regulation and consequences of protein SUMOylation. Biochem. J. **428**, 133–145
46 Miller, M. J., Barrett-Wilt, G. A., Hua, Z. and Vierstra, R. D. (2010) Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in *Arabidopsis*. Proc. Natl. Acad. Sci. U.S.A. **107**, 16512–16517
47 Budhiraja, R., Hermkes, R., Muller, S., Schmidt, J., Colby, T., Panigrahi, K., Coupland, G. and Bachmair, A. (2009) Substrates related to chromatin and to RNA-dependent processes are modified by *Arabidopsis* SUMO isoforms that differ in a conserved residue with influence on desumoylation. Plant Physiol. **149**, 1529–1540
48 Elrouby, N. and Coupland, G. (2010) Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify *Arabidopsis* proteins implicated in diverse biological processes. Proc. Natl. Acad. Sci. U.S.A. **107**, 17415–17420

Received 15 September 2010/10 March 2011; accepted 17 March 2011
Published as BJ Immediate Publication 17 March 2011, doi:10.1042/BJ20101446

Biochem. J. (2011) **436**, 581–590 (Printed in Great Britain) doi:10.1042/BJ20101446

online data

SUPPLEMENTARY ONLINE DATA

Distinctive properties of *Arabidopsis* SUMO paralogues support the *in vivo* predominant role of AtSUMO1/2 isoforms

Laura CASTAÑO-MIQUEL, Josep SEGUÍ and L. Maria LOIS[1]

Department of Molecular Genetics, Center for Research in Agricultural Genomics (CRAG) CSIC-IRTA-UAB, Parc de Recerca UAB, Edifici CRAG, Campus UAB, 08193 Bellaterra, Barcelona, Spain

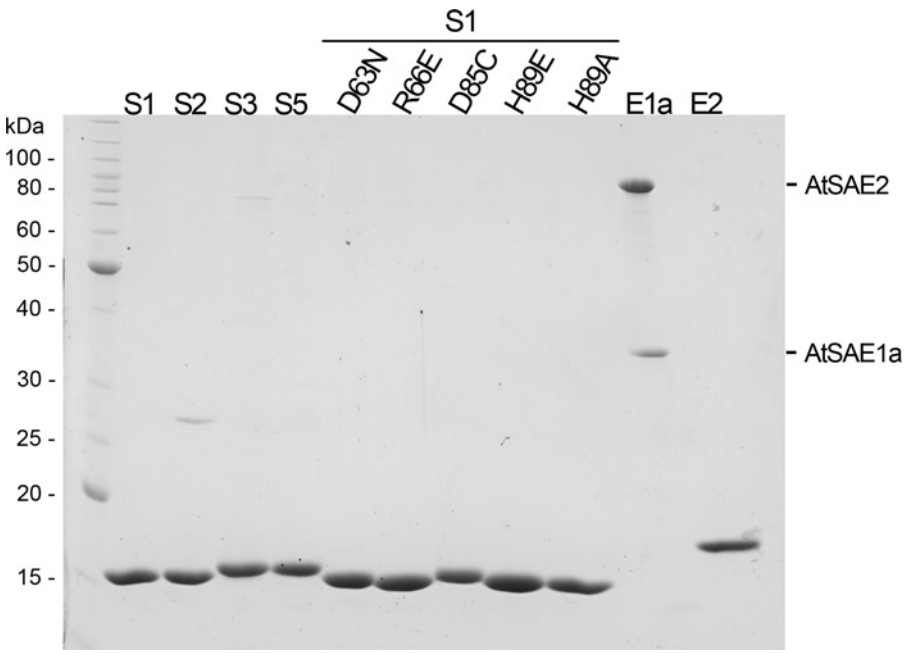


Figure S1 Purification from *E. coli* (DE3) BL21 of the SUMO conjugation machinery components

Purified protein sample (2 μg of each) was analysed SDS/PAGE (12 % gel) and stained with Coomassie Blue.

[1] To whom correspondence should be addressed (email maria.lois@cid.csic.es).

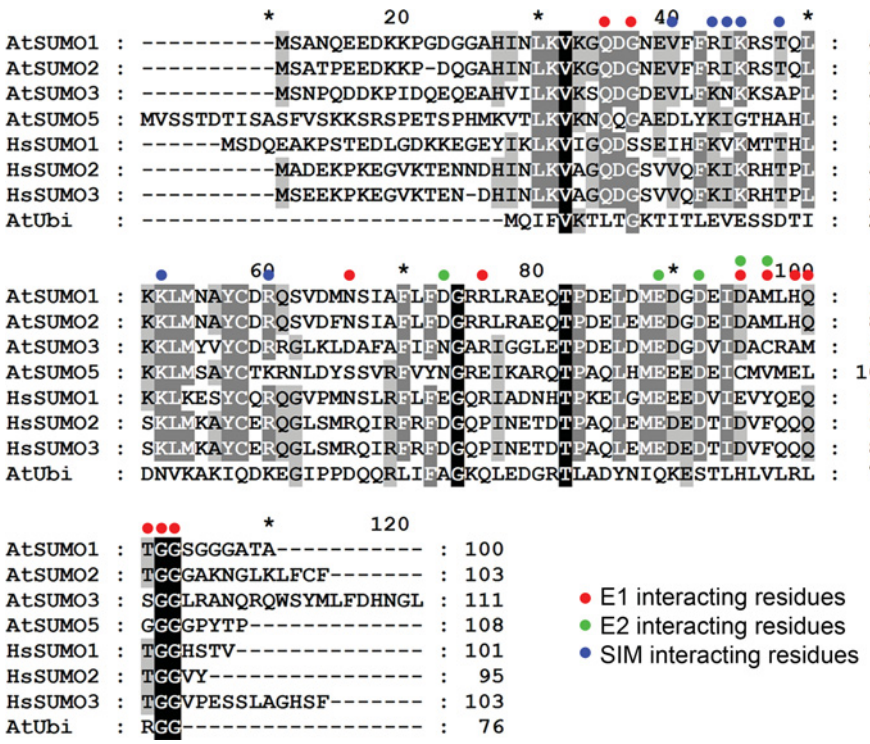


Figure S2 Sequence alignment of *Arabidopsis* (At) and human (Hs) SUMO orthologues and *Arabidopsis* ubiquitin (AtUbi)

Dots above the alignment indicate residues involved in E1 contacts (red), E2 non-covalent interactions (green) and interaction with SIMs (blue).

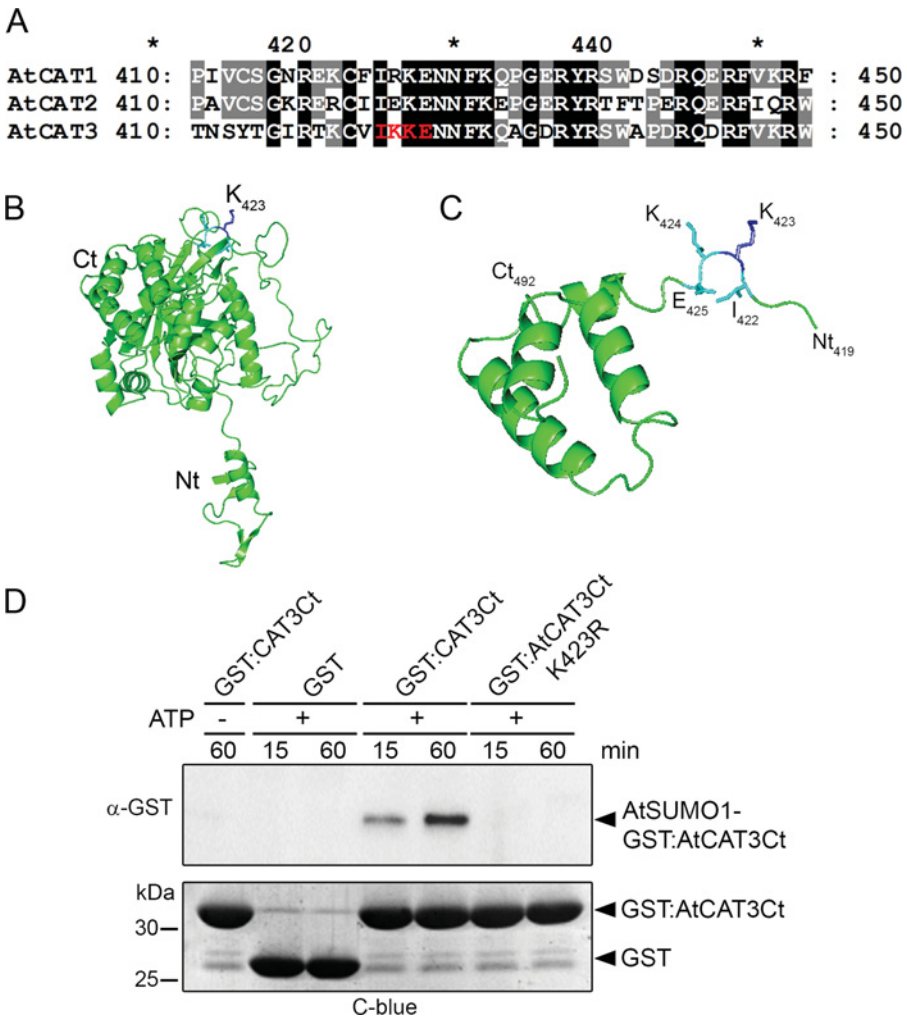


Figure S3 ***In vitro*-specific SUMOylation of the AtCAT3 C-terminal domain**

(**A**) Protein sequence alignment of the three *Arabidopsis* catalase isoform regions containing the predicted SUMOylation consensus site, which is shown in red. (**B**) AtCAT3 structure as predicted by the SWISS-MODEL comparative protein modelling server (shown as a green ribbon). *Exiguobacterium oxidotolerans* catalase structure was used as a template (PDB code 2J2M). (**C**) Structure model of the AtCAT3 C-terminal domain (comprising amino acids 419–492) used as a GST-fusion protein in the *in vitro* SUMOylation assays. A lateral chain of amino acids in the SUMOylation consensus site is represented by sticks: dark blue for the acceptor lysine (Lys^{423}) and light blue for the others (Iso^{422}, Lys^{424} and Glu^{425}). (**D**) *In vitro* SUMOylation assays were performed in the presence of AtE1a, AtSUMO2, AtSCE1 and the corresponding substrate. Incubation was performed at 227 °C and stopped at the specified time points. Reaction products were resolved by SDS/PAGE and examined by immunoblot analysis with anti-GST antibodies. The portion of the Coomassie Blue (C-Blue)-stained membrane that corresponds with the GST substrates is shown as a loading control.

Received 15 September 2010/10 March 2011; accepted 17 March 2011
Published as BJ Immediate Publication 17 March 2011, doi:10.1042/BJ20101446

CAPÍTULO-II

Diversification of SUMO-Activating Enzyme in *Arabidopsis:* Implications in SUMO conjugation

Diversification of SUMO-Activating Enzyme in *Arabidopsis*: Implications in SUMO conjugation.

Castaño-Miquel L[a,2], Seguí J[a,2], Manrique S[a,2], Teixeira I[a], Carretero-Paulet L[b], Atencio F[a] and Lois LM[a]

a Center for Research in Agricultural Genomics CRAG (CSIC-IRTA-UAB), Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.
b Institute for Plant Molecular and Cell Biology-IBMCP (CSIC-UPV) Integrative Systems Biology Group C/ IngenieroFaustoElio s/n. 46022 Valencia, Spain.
2 These authors contributed equally to this work.

Mol Plant. 2013 Sep; 6(5):1646-60. doi: 10.1093/mp/sst049. Epub 2013 Mar 12.

PMID: 23482370 [PubMed – indexed for MEDLINE]

Resumen. SUMO, Small Ubiquitin-likeMOdifier, es un modificador post-traduccional de la familia de la Ubiquitina. La sumoilación está finamente regulada y la unión final de SUMO al substrato, modula diferentes procesos biológicos esenciales en plantas, sin embargo, los mecanismos moleculares que regulan todas estas funciones biológicas de SUMO aún se desconocen.

En este trabajo se han evaluado las implicaciones bioquímicas y biológicas de la enzima heterodimérica E1 de SUMO, que cataliza el primer paso de la cascada enzimática de SUMO. Se ha establecido que la subunidad grande, SAE2, y la subunidad pequeña, SAE1, de *A.thaliana* están codificadas por uno y tres genes, respectivamente. Los tres genes parálogos *SAE1a*, *SAE1b1* y *SAE1b2* son el resultado de dos eventos de duplicaciones independientes. *In vivo*, solamente existen dos isoformas de la enzima activadora, E1a y E1b, debido a que los genes *SAE1b*1 y *SAE1b2* corresponden a dos copias idénticas duplicadas en tándem en el genoma de *Arabidopsis*. Ambos heterodímeros muestran una localización nuclear, dependiente de una señal de localización nuclear situada en el extremo C-terminal de SAE2. Se ha determinado que la cinética de conjugación de SUMO *in vitro* es dependiente de la isoforma de la enzima E1 presente en el ensayo. Además, estudios *in vivo* muestran que los mutantes *atsae1a* presentan defectos de SUMOilación en respuesta a estrés térmico e hídrico, consistente con el fenotipo que presentan plantas con la SUMOilación afectada. Nuestros resultados apoyan estudios anteriores que apuntan a un papel regulador de la enzima activadora de SUMO, E1, durante la conjugación de SUMO *in vivo* y proporcionan un nuevo

mecanismo de regulación de la SUMOilación en las plantas mediante la existencia de variantes de la enzima E1 activadora de SUMO.

Diversification of SUMO-Activating Enzyme in *Arabidopsis*: Implications in SUMO Conjugation

Laura Castaño-Miquel[a,2], Josep Seguí[a,2], Silvia Manrique[a,2], Inês Teixeira[a], Lorenzo Carretero-Paulet[b], Félix Atencio[a] and **L. Maria Lois[a,1]**

a Center for Research in Agricultural Genomics CRAG (CSIC-IRTA-UAB-UB), Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain
b Department of Biological Sciences, SUNY-University at Buffalo, North Campus, 109 Cooke Hall, Buffalo, NY 14260, USA

ABSTRACT Sumoylation is an essential posttranslational modification that participates in many biological processes including stress responses. However, little is known about the mechanisms that control Small Ubiquitin-like MOdifier (SUMO) conjugation *in vivo*. We have evaluated the regulatory role of the heterodimeric E1 activating enzyme, which catalyzes the first step in SUMO conjugation. We have established that the E1 large SAE2 and small SAE1 subunits are encoded by one and three genes, respectively, in the *Arabidopsis* genome. The three paralogs genes *SAE1a*, *SAE1b1*, and *SAE1b2* are the result of two independent duplication events. Since *SAE1b1* and *SAE1b2* correspond to two identical copies, only two E1 small subunit isoforms are present *in vivo*: SAE1a and SAE1b. The E1 heterodimer nuclear localization is modulated by the C-terminal tail of the SAE2 subunit. *In vitro*, SUMO conjugation rate is dependent on the SAE1 isoform contained in the E1 holoenzyme and our results suggest that downstream steps to SUMO–E1 thioester bond formation are affected. *In vivo*, SAE1a isoform deletion in T-DNA insertion mutant plants conferred sumoylation defects upon abiotic stress, consistent with a sumoylation defective phenotype. Our results support previous data pointing to a regulatory role of the E1 activating enzyme during SUMO conjugation and provide a novel mechanism to control sumoylation *in vivo* by diversification of the E1 small subunit.

Key words: E1 activating enzyme; conjugation rate; subcellular localization; regulation; abiotic stress; gene duplication.

INTRODUCTION

In eukaryotic cells, posttranslational modifications by SUMO (Small Ubiquitin-like MOdifier) modulate protein activity through regulation of subcellular localization, protein activity and stability, and protein–protein interactions (Wilkinson and Henley, 2010). SUMO is synthesized as a precursor that is processed via specific proteases, ULP, releasing a SUMO mature form with a Gly–Gly motif at its C-terminus. SUMO is conjugated to protein targets through three sequential reactions. In a first step, SUMO is activated by the heterodimeric E1-activating enzyme in an ATP-dependent reaction. Activated SUMO is transferred to the E2-conjugating enzyme as a prior step to conjugation to the substrate, which can be achieved by direct transfer from the E2 or facilitated by E3 ligases. Sumoylation is a reversible modification and SUMO excision from the substrate is catalyzed by the same class of cysteine proteases involved in the maturation step (Gareau and Lima, 2010).

Many studies have addressed the biological function of SUMO in plants by using the dicot *Arabidopsis* and the monocot rice as models. In both models, SUMO has a role in response to abiotic stresses, such as heat, cold, salt, and ABA signaling, in addition to a wide array of developmental processes (Chaikam and Karlson, 2010; Lois, 2010; Miura and Hasegawa, 2010; Park et al., 2010; Thangasamy et al., 2011; Wang et al., 2011). In rice, sumoylation has also been involved in hybrid male sterility (Long et al., 2008). Moreover, defense responses (Kim et al., 2008), flowering (Murtas et al., 2003),

[1] To whom correspondence should be addressed. E-mail maria.lois@cragenomica.es, tel. +34 93 5636600 ext. 3215, fax +34 93 5636601.

2 These authors contributed equally to this work.

doi: 10.1093/mp/sst049, Advance Access publication 12 March 2013

Received 14 February 2013; accepted 3 March 2013

nitrogen metabolism (Park et al., 2011), phosphate starvation (Miura et al., 2005), drought tolerance (Catala et al., 2007), and sensitivity to copper (Chen et al., 2011) are processes affected by altered sumoylation in *Arabidopsis*.

Most of the mentioned studies used plants with mutations in E3 ligases and ULP proteases. Among them, the most studied mutants are the *siz1* E3 ligase and the *esd4* ULP protease. *siz1*-null mutant displays a reduction in endogenous SUMO conjugate accumulation while an overaccumulation of SUMO conjugates is found in *esd4* mutant. Even though they have opposite molecular effects, the physiological outcome of these mutations is very similar. An explanation to these observations is that the molecular mechanisms that mediate the severe growth defects found in those mutants might be different (Hermkes et al., 2011). More importantly, these results indicate that SUMO conjugation homeostasis is under a tight control and over- or under-accumulation of SUMO conjugates results in a misregulation of essential processes.

In the sumoylation cascade, the role of the E1 activating enzyme, which is the first control point in the conjugation pathway, is particularly interesting. The E1 is a heterodimeric enzyme consisting of a large subunit, SAE2, and a small subunit, SAE1. SAE2 is structured in four functional domains: adenylation, catalytic cysteine, ubiquitin-fold (UFD), and C-terminal domains (Lois and Lima, 2005). On the other hand, SAE1 contributes the essential Arg21 to the adenylation domain (Lee and Schindelin, 2008). The adenylation domain is responsible for SUMO recognition and SUMO C-terminus adenylation, as a prior step to thioester bond formation with the E1 catalytic cysteine. By using SUMO intermediate analogs, structural studies have established that thioester bond formation requires an active site remodeling that involves a 130 degree rotation of the catalytic cysteine domain (Olsen et al., 2010).

In contrast, much less is known about how E1 is regulated *in vivo*. In mammalian cells, SUMO E1 is localized to the nucleus by a nuclear localization signal located in the C-terminal tail of the large subunit SAE2 (Moutty et al., 2011). Also, low ROS levels result in SUMO conjugation inhibition by inducing the formation of a disulphide bridge between the catalytic Cys residues of the E1 activating and the E2 conjugating enzymes. This inhibition is reversible and could constitute an early step in ROS signaling (Bossis and Melchior, 2006).

In plants, the regulation of SUMO conjugation is far from being understood. The SUMO system is apparently more complex in terms of number and the biochemical properties of its components. In *Arabidopsis*, SUMO1 and 2 have diverged to the extent that double mutants *sumo1sumo2* cannot be complemented by SUMO3 and 5 (Saracco et al., 2007). This divergence is reflected by a preferential conjugation of SUMO1/2 versus SUMO3 and 5 through a mechanism involving SUMO isoform selection by the E1 (Castaño-Miquel et al., 2011). In addition to SUMO, other components of the SUMO conjugation system have also diverged into several isoforms, such as the E1 and E2 (Novatchkova et al., 2012). It has been proposed that *Arabidopsis* expresses two isoforms of the E1 activating enzyme that differ in the small subunit composition, SAE1a or SAE1b, whereas the large subunit, SAE2, is unique. The existence of a functional specialization of both isoforms has not been reported so far.

Here, we have analyzed the biological implications of E1 activating enzyme diversification. We showed that the E1 activating enzyme is encoded by one variant of the *SAE2* gene and three copies of the *SAE1* gene: *SAE1a*, *SAE1b1*, and *SAE1b2*. Since *SAE1b1* and *SAE1b2* are identical, only two isoforms of the E1 activating enzyme can be found: SAE2/SAE1a (E1a) and SAE2/SAE1b (E1b). The heterodimeric E1 subcellular localization is facilitated by the C-terminal tail of the SAE2 subunit. Even though most of the E1 functional domains are located in the SAE2 subunit, we found that SAE1a and SAE1b isoforms conferred distinct conjugation rates *in vitro*, suggesting a regulatory role of the SAE1 small subunit during SUMO conjugation. *sae1a* mutant plants, which do not express SAE1a isoform, displayed defects in SUMO conjugation upon heat and drought responses, which is characteristic of SUMOylation-deficient plants (Catala et al., 2007; Saracco et al., 2007). These results indicate that SAE1a is necessary for maintaining SUMO conjugation homeostasis *in vivo* and that SAE1a and SAE1b are not fully redundant activities. Diversification of the E1 enzyme is not restricted to *Arabidopsis*, suggesting that the presence of several E1 isoforms might constitute an evolutionary advantage. Taken together, we postulate that the E1 small subunit SAE1 could be modulating downstream steps during SUMO conjugation providing a novel mechanism to control sumoylation *in vivo*.

RESULTS

Genes Encoding the SUMO-Activating Enzyme

According to annotations in the Arabidopsis Information Resource (TAIR), two variants of the gene encoding the E1 large subunit SAE2 could exist (Figure 1A). The second variant would result from an alternative splicing that would generate a mature mRNA retaining the last intron. In order to identify this second variant, we extracted total RNA from several tissues and used it as a template for cDNA synthesis using oligo dT. The resulting cDNA was analyzed by PCR using three primer pair combinations (Figure 1A). We only detected PCR products when primers annealing to the *SAE2 variant 1* were present in the reaction. These products were unique and displayed an Rf consistent with the presence of *SAE2 variant 1* (SM020–SM004 and SM019–SM004). When primers annealing specifically to *SAE2 variant 2* were used (SM021–SM004), no reaction products were obtained in any of the analyzed samples, suggesting that *SAE2 variant 2* does not exist *in vivo* and that is the result of an annotation error (Figure 1B). Moreover, three gene accessions codifying for the E1 small subunit SAE1 are found: At4g24940, At5g50580, and At5g50680. The first accession codifies for the SAE1 isoform a, and the others correspond to two exact copies of a gene codifying for the SAE1 isoform b located in tandem

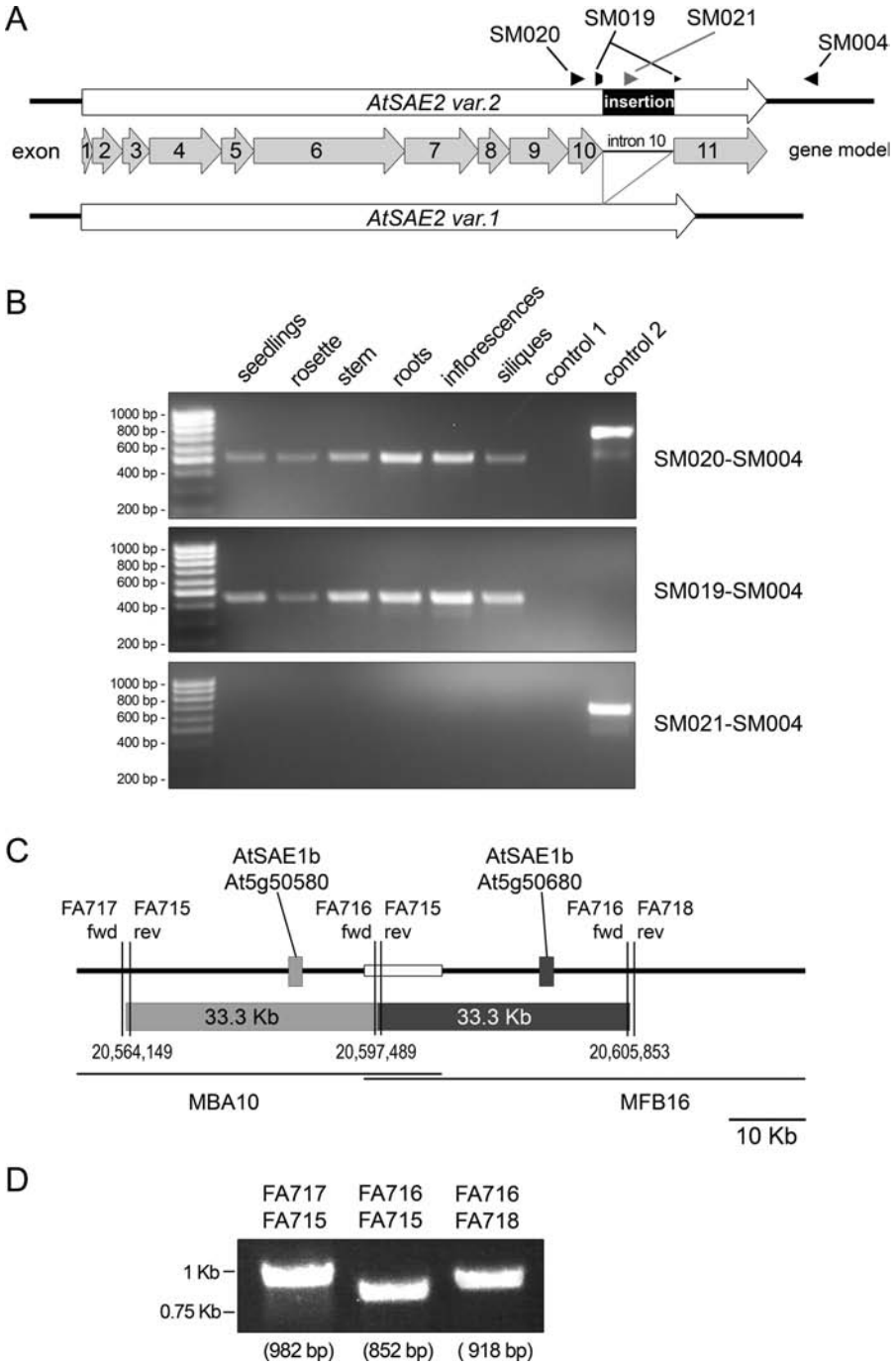


Figure 1. Gene Structure of the *Arabidopsis* E1-Activating Enzyme.

(A) Schematic representation of the *SAE2* variant 2 as annotated in Arabidopsis Information Resource (TAIR) is shown on top. The insertion in *SAE2* variant 2 corresponds to intron 10 (middle) that is excised in variant 1 (bottom). The annealing sites of the primers used to analyze the *SAE2* variant 2 expression are shown above the scheme.

(B) RNA extracted from the specified tissues was retrotranscribed using oligo dT and the resulting cDNA synthesized was analyzed by PCR using the primers pair indicated on the right. PCR products were resolved on agarose gel electrophoresis and stained with ethidium bromide. Control 1 corresponds to RT negative control and control 2 contains Col0 genomic DNA as a template.

(C) Schematic representation of the *Arabidopsis* chromosome 5 genomic region containing a 33-Kb-sequence duplication. SAE1b genomic regions are represented by light gray and dark gray boxes. The MBA10 and MFB16 overlapping region is indicated by a white rectangle (top). Duplicated regions sharing a 99% of DNA sequence identity are represented by light gray and dark gray rectangles, and their coordinates in chromosome 5 are also indicated (middle). Contigs containing the represented genomic region are indicated at the bottom, MBA10 and MFB16.

(D) PCR products from reactions performed in the presence of the indicated primers were analyzed as in (B). The expected size of each PCR product is indicated below each lane.

(Figure 1C). PCR analysis of *Arabidopsis* genomic DNA confirmed the existence of a 33.3-Kb tandem duplication that includes the gene codifying for SAE1b (Figure 1D). Sequencing of the PCR product synthesized using the primer pair FA716–FA715 confirmed that the amplified fragment corresponded to the region connecting both tandem repeats. This genomic duplication affects 11 additional genes (Supplemental Table 1). Comparative genome analysis indicated that this duplication is not present in the *Arabidopsis thaliana* ecotypes Ler-1, C24, Bur-0, and Kro-0 (Supplemental Figure 1).

Evolutionary Diversification of SUMO-Activating Enzyme among Plants

We next examined SUMO E1 evolutionary diversification in plants. Search for *SAE1* genes resulted in the identification of 39 sequences, including the SAE1 family from 31 species with fully sequenced genomes plus a gymnosperm representative from *Picea sitchensis* (Supplemental Table 2). Similarly, searches for *SAE2* genes returned 33 sequences corresponding to the SAE2 family from 27 species with fully sequenced genomes (Supplemental Table 3). Both sets of sequences included representative of the main plant evolutionary lineages. To investigate the evolutionary relationships among *SAE2* and *SAE1* genes in plants, we performed phylogenetic analyses on the basis of protein sequence alignments. Three alternative methods of phylogenetic reconstruction were used: Bayesian inference, maximum likelihood, and neighbor joining. Bayesian phylogenetic trees are shown in Figure 2. Trees obtained from the three alternative methods show almost identical topologies, except for a few internal nodes. The exon/intron structures of *SAE1* and *SAE2* genes are also displayed next to the corresponding trees. The intron positions and exon phases are remarkably well conserved, providing further support to the phylogenetic analyses. Only minor differences could be observed in intron number and length.

In both cases, clustering of the sequences in the trees reflects quite well the taxonomical relationships of the species represented. Two independent clades grouping dicot and monocot SAE1 and SAE2 sequences were retrieved with high statistical support. Many plant species show SAE1 and SAE2 families composed of a single gene. However, phylogenetic analyses reveal lineage-specific gene duplication within the SAE1 and SAE2 families. This is the case of *Arabidopsis thaliana*, displaying three *SAE1* genes resulting from two duplication events. One likely resulted from a polyploidization event predating the emergence of the Brassicaceae lineage. Consistently, *A. thaliana At5G50580* and *At4G24940* genes map in genomic regions of chromosomes 5 and 4 resulting from a whole-genome duplication event estimated to have occurred in the last 25–40 million years (Blanc et al., 2003), matching the time of origin and diversification of the Brassicaceae family (Couvreur et al., 2010). A second involved a more recent tandem duplication likely occurring after divergence of *A. thaliana* Col-0 from other ecotypes (Supplemental Figure 1). Besides *A. thaliana* Col-0, the remaining Brassicaceae species display two *SAE1* genes, with the exception of only *Brassica rapa*, which displays a single *SAE1* gene. Interestingly, searches of the *B. rapa* genome identify two additional sequences with shorter lengths showing similarity with SAE1 (Bra023205 and Bra03552), maybe corresponding to *SAE1* pseudogenes. Additional lineage-specific *SAE1* gene duplications could be observed in *Glycine max*, *Populus trichocarpa*, and *Manihot esculenta*. These three species, together with *Malus domestica*, *B. Rapa*, and the basal land plant species *Physcomitrella patens* also show two *SAE2* genes resulting from gene duplication events occurring recently in the corresponding lineages.

Conservation and Structure Prediction of SUMO-Activating Enzyme Isoforms

In yeast and mammals, there is a single form of the SUMO E1 enzyme. In contrast, *Arabidopsis* expresses two forms of the E1 that differ in the small subunit composition. The *Arabidopsis* E1 large subunit SAE2 shares a 36% of amino acid sequence identity with its human ortholog HsSAE2, although this conservation is not evenly distributed between functional domains. The adenylation domain (residues 1–153 + 379–432) is the most conserved, with a 55% of sequence identity, the C-terminal tail is the less conserved region (16% identity; residues 546–625), and the catalytic cysteine (residues 154–378) and the UFD domains (residues 433–545) share a 31% of sequence identity with their human counterparts (Figure 3A and 3B). Similarly, when we analyzed the conservation degree among plant SAE2 paralogs, the SAE2 adenylation domain and the C-terminal tail presented the highest and the lowest conservation degree, respectively.

The two *Arabidopsis* isoforms of the E1 activating enzyme small subunit, SAE1a and SAE1b, are highly conserved. The analysis of the E1 structure model showed that the SAE1a/b regions involved in the contacts with the large subunit have a larger conservation degree with its human ortholog (41–44% of sequence identity; residues 6–106 + 271–317), whereas, sequences located in distant regions from the heterodimer interface are more divergent (21–20% of sequence identity; residues 107–270; Supplemental Figure 3). The most conserved regions correspond to those domains in the human SAE1 ortholog that undergo conformational changes and/or that contain residues that establish different contacts between the E1 open and closed conformations during SUMO activation. These residues are also conserved between *Arabidopsis* SAE1 paralogs. On the other hand, we could not identify any domain susceptible to determine functional specialization between SAE1a and SAE1b since non-conserved residues are evenly distributed among these paralogs (Figure 3C).

In Vitro SUMO Conjugation Rate Is Dependent on SUMO-Activating Enzyme Isoforms

Arabidopsis SAE1 paralogs display a 20% of sequence divergence and we analyzed a possible effect of these differences

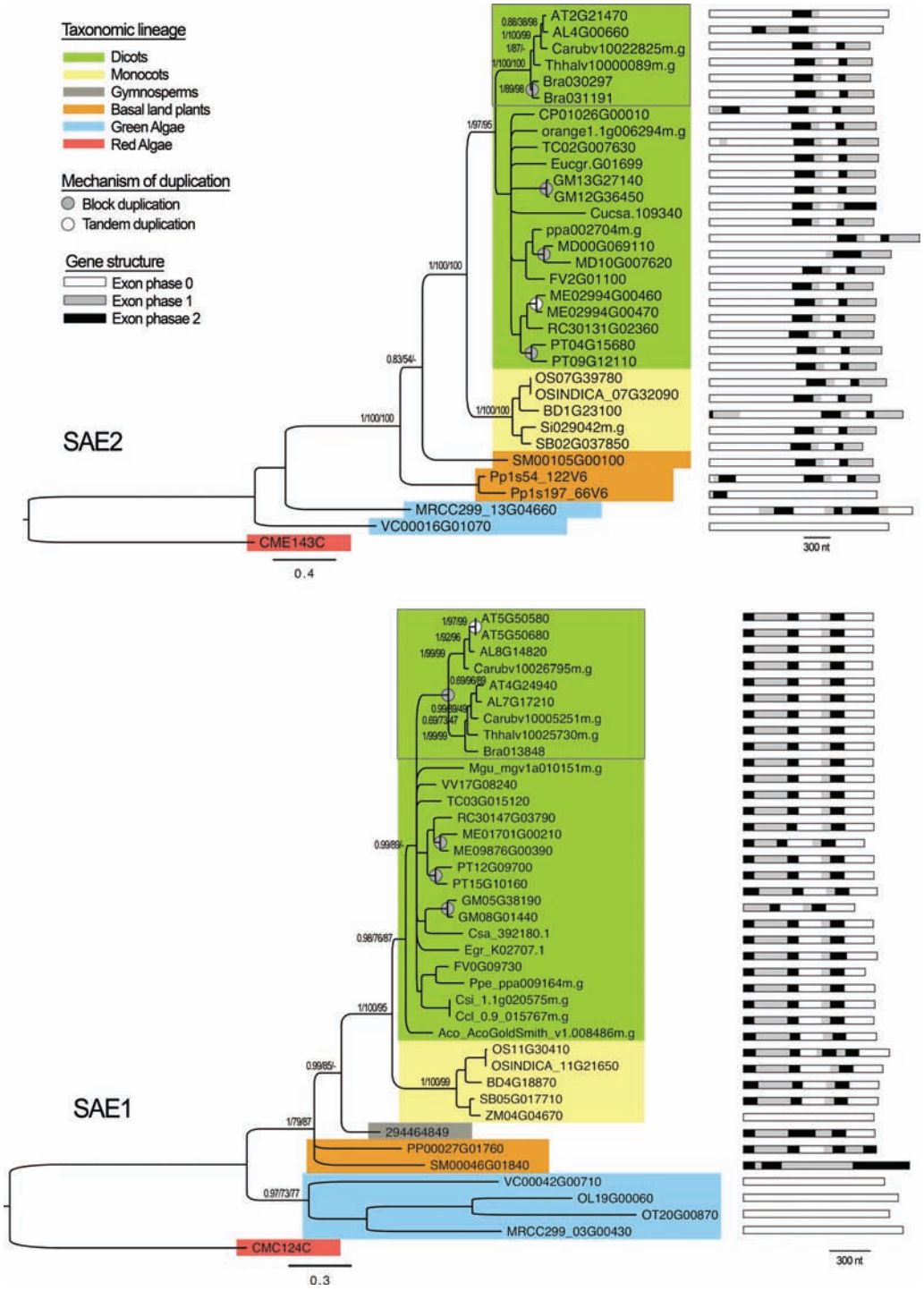


Figure 2. Bayesian Phylogeny and Exon–Intron Structure of Plant SUMO-Activating Enzyme (*SAE*) Genes.
Bayesian phylogenetic tree depicting the evolutionary relationships among 39 *SAE1* protein sequences from 32 plant species **(A)**, and 33 *SAE2* protein sequences from 27 plant species **(B)**. Both trees were rooted using the corresponding ortholog from the red algae *Cyanidioschyzon merolae*. Values next to the nodes indicate statistical support on relevant clades (Bayesian posterior probabilities/maximum-likelihood a LRT support values/neighbor-joining bootstrap values). The tree is drawn to scale, with branch lengths proportional to evolutionary distances between nodes. The scale bar indicates the estimated number of amino acid substitutions per site. The origin of gene duplication (tandem versus block duplication) is indicated at the corresponding nodes. Clades clustering Brassicaceae representatives are highlighted within a dashed box. Exon organization is shown on the right, with boxes colored according to phases, except for the *P. sitchensis* SAE1 representative, for which the genomic sequence was not available.

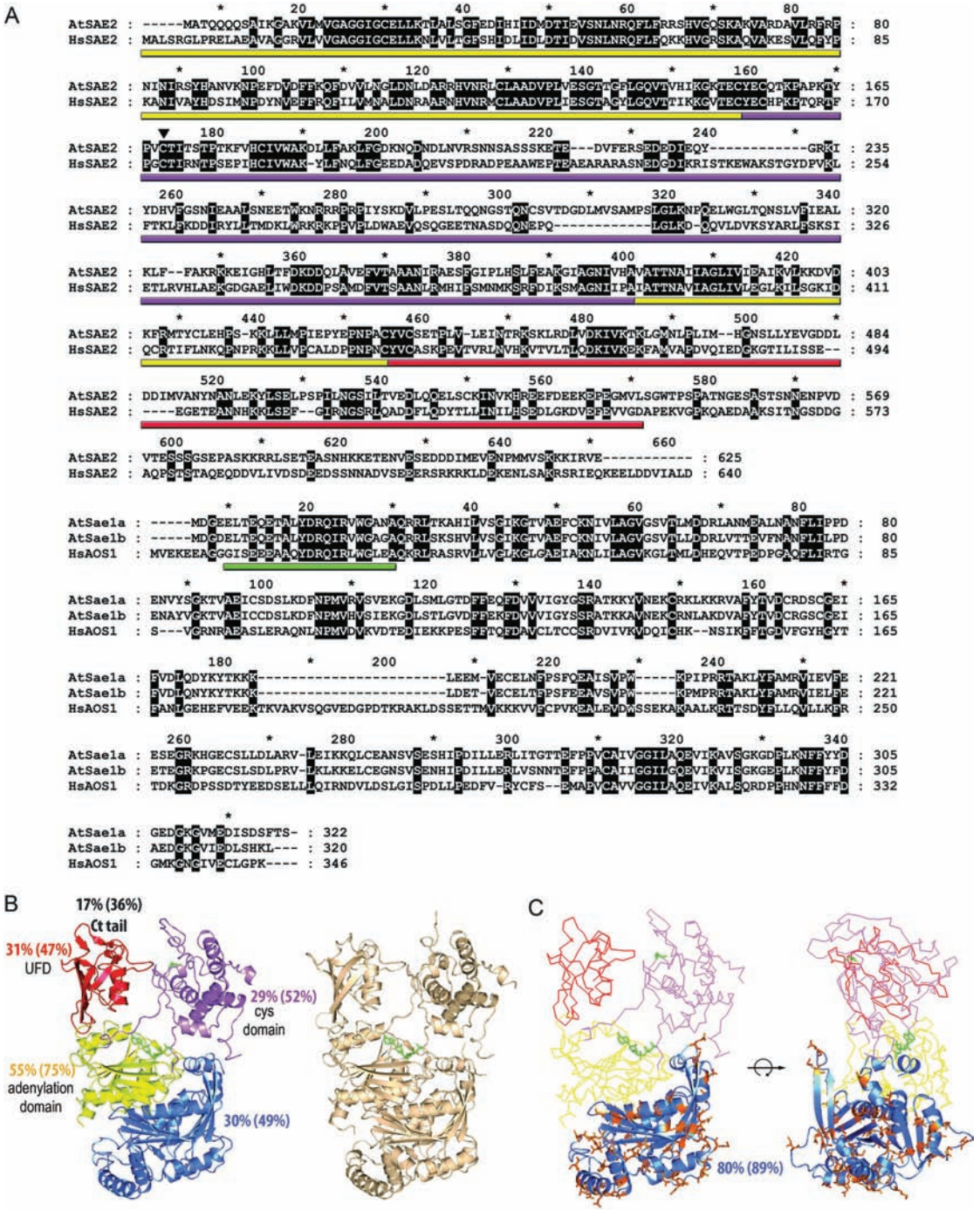


Figure 3. SUMO-Activating Enzyme Isoforms Conservation.
(A) Amino acid sequence alignment of *Arabidopsis* (At) and human (Hs) orthologs of E1 SUMO-activating enzyme large subunit (SAE2; top) and small subunit (SAE1; bottom). Black background and white letters correspond to 100% sequence identity. Distribution of SAE2 functional domains, adenylation domain (yellow; 1–153 + 379–432), cysteine domain (violet; 154–378), and the ubiquitin-fold domain (UFD, red; 433–545) is shown below the sequence. The C-terminal tail (546–625) is not underlined. The catalytic cysteine residue is indicated by an arrowhead. The SAE1 region containing residues that undergo conformational changes during SUMO activation is underlined (green).
(B) *Arabidopsis* SUMO E1 structure, as predicted by the SWISS-MODEL comparative protein modeling server, and the human SUMO E1 structure used as a template (1Y8Q) are shown as a ribbon diagram on the right and the left, respectively. SAE2 functional domains are colored as in (A), the catalytic cysteine and ATP are in green, and the SAE1 subunit is in blue. Sequence identity percentage is shown by domain and similarity is indicated in parentheses. The *Arabidopsis* SAE2 C-terminal tail is not represented, since is not resolved in the original structure template.
(C) Non-conserved amino acids between SAE1a and SAE1b paralogs are shown in brown and side chains are represented as sticks. Sequence identity and similarity are shown as in (B).

in the heterodimer activity. We tested SUMO conjugation efficiency in reactions containing the SUMO-activating enzyme isoform a (E1a: SAE2/SAE1a) or isoform b (E1b: SAE2/SAE1b). Time-course sumoylation reactions incubated at 22°C, 37°C, and 42°C showed that SUMO conjugation increased with temperature (Figure 4A–4C). Quantification of relative reaction efficiency among E1a and E1b isoforms indicated that reactions containing the E1a isoform conjugated SUMO significantly more efficiently at 37°C and 42°C, although these differences were not significant at 22°C (Figure 4E). We investigated whether these differences were related to SUMO–E1 thioester formation. Our results showed that the E1 holoenzyme containing the isoform SAE1a was also slightly more efficient at establishing SUMO–E1 thioester bonds but this effect was restricted to high temperatures (42°C and 48°C) (Figure 4D and 4F). These results suggest that E1 small subunit isoform could have an effect in downstream steps to SUMO–E1 thioester bond formation.

Subcellular Localization of SUMO-Activating Enzyme Isoforms

SAE2 and SAE1a/b amino acid sequences were analyzed using the software for prediction of protein localization PSORT (http://wolfpsort.org/). As a result, five putative nuclear localization signals were identified in SAE2 and none in SAE1a/b. Four of them belong to the NLS pat4 type, located at the amino acid positions 257, 258, 326, and 582, and the fifth belongs to the NLS pat7 type and it is located at the amino acid position 579 (Figure 5A). To perform the functional analysis of the predicted NLS motifs, the enhanced yellow fluorescent protein (EYFP) was fused to the C-terminus of full-length SAE2 (residues 1–625) or the C-terminal deletion mutant SAE2ΔCt

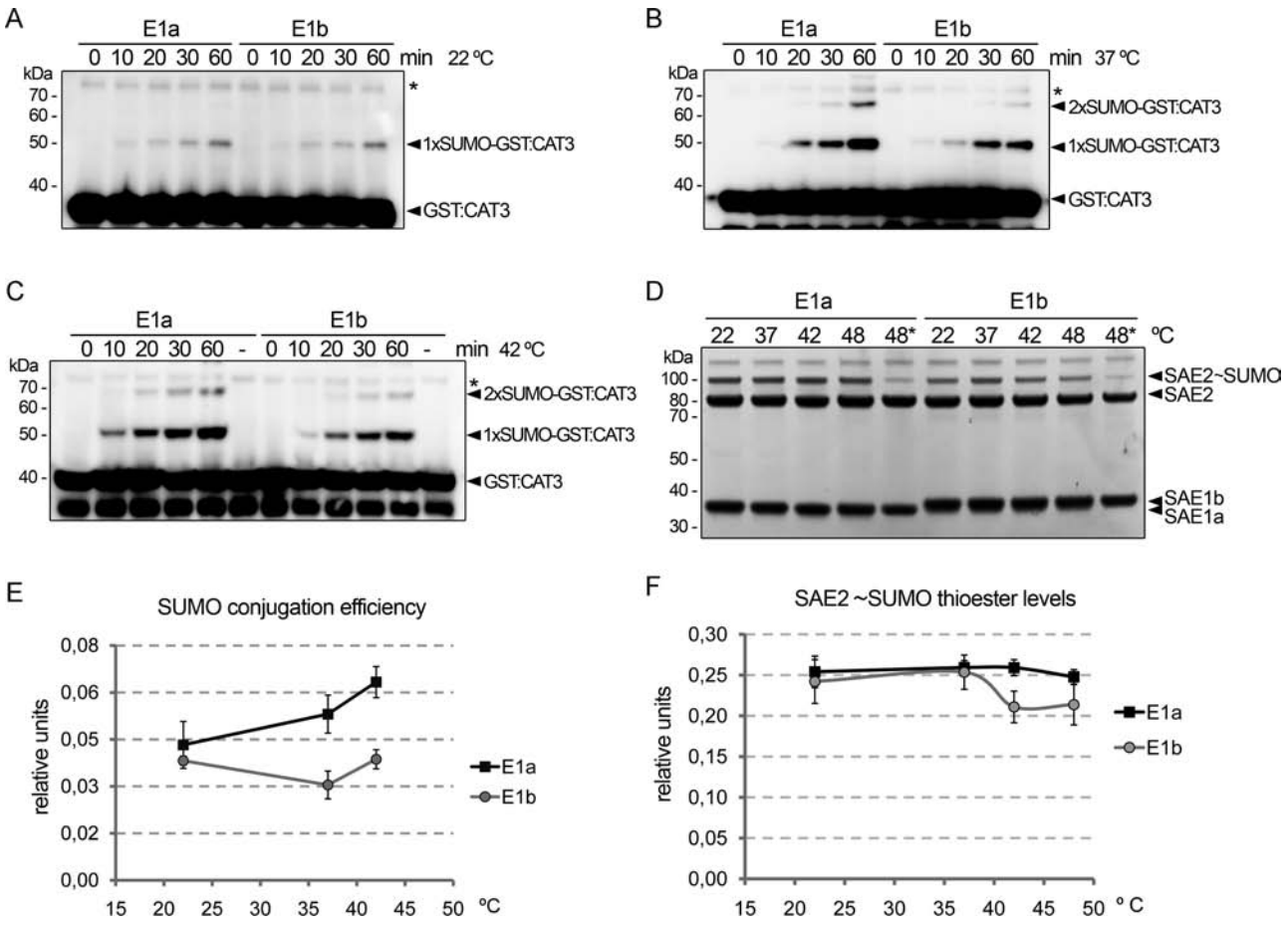


Figure 4. *In Vitro* Characterization of *Arabidopsis* SUMO-Activating Enzyme Isoforms.
(A–C) *In vitro* sumoylation assays were performed at 22°C, 37°C, and 42°C in the presence of E1a or E1b, SUMO2, SCE1, and GST:CAT3Ct as a substrate. Reaction mixtures were stopped at the indicated time, and products were resolved by SDS–PAGE and examined by immunoblot analysis with anti-GST antibodies. The asterisk indicates a contaminating protein.
(D) SAE2~SUMO thioester formation assays were performed using E1a (SAE2:SAE1a) or E1b (SAE2:SAE1b) and SUMO1. Reaction mixtures were incubated at the indicated temperatures and stopped after 2 min. As a control for thioester bond formation, an aliquot of the reaction incubated at 48°C was treated with DTT (48*). Reaction products were separated by SDS–PAGE and stained with Coomassie fluor orange.
(E, F) Reactions were performed at least in triplicate and GST:CAT3Ct sumoylation efficiency or SAE2~SUMO thioester bond formation level quantified. Average values and standard deviation bars were plotted on the graphs.

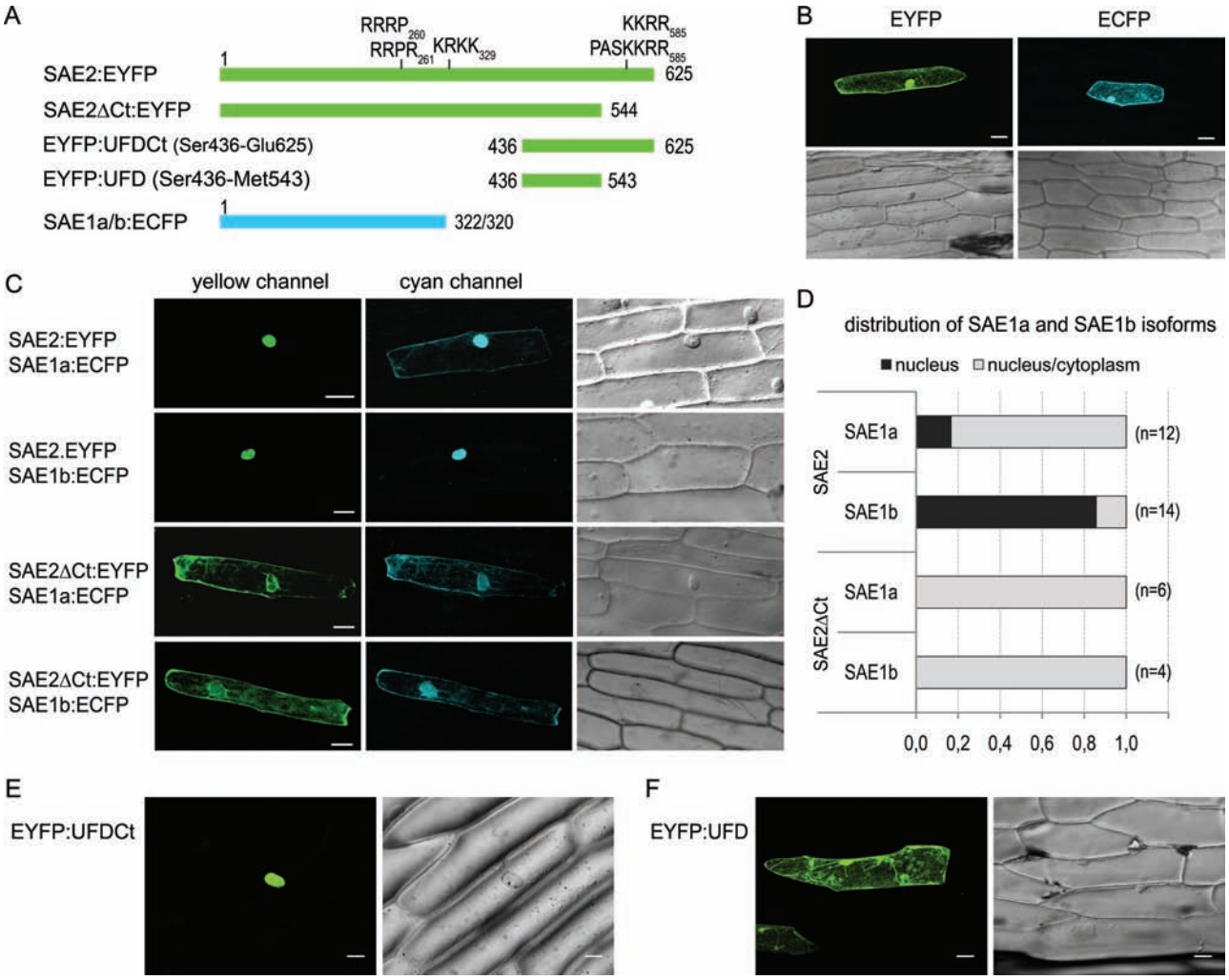


Figure 5. Subcellular Localization of *Arabidopsis* SUMO E1-Activating Enzyme Isoforms.
(A) Full-length SAE2, SAE2 with a C-terminal tail deletion, and SAE2 domain S436–E625 were fused to the EYFP (enhanced yellow fluorescent protein) and transiently expressed in onion epidermal cells. For co-localization studies, SAE1a/b were fused to the ECFP (enhanced cyan fluorescent protein). Nuclear localization signals as predicted by the PSORT software (http://psort.ims.u-tokyo.ac.jp/) are shown above the SAE2 representation and their position on the protein sequence indicated.
(B) Cells expressing EYFP or ECFP were used as control.
(C) E1a and E1b localization experiments were performed by co-expression of SAE2:EYFP and SAE1a:ECFP, SAE2ΔCt:EYFP and SAE1a:ECFP, and EYFP:SAE2 and SAE1b:ECFP. Light transmission images of the onion epidermal cells are also shown. Bars = 50 μm.
(D) Subcellular distribution of SAE2, SAE2ΔCt, SAE1a, and SAE1b was scored and represented on the plot (*n* indicates the number of cells analyzed).
(E) EYFP:SAE2 Ser436–Glu625 and **(F)** EYFP:SAE2 Ser436–Met543 localization. Bars = 50 μm.

(residues 1–545), which retains the three predicted NLS motifs located at positions 257, 258, and 326 but not the two motifs located at positions 579 and 582. The resulting fusion proteins are indicated as SAE2:EYFP and SAE2ΔCt:EYFP, respectively. In addition, enhanced cyan fluorescent protein (ECFP) was fused to the C-terminus of SAE1a or SAE1b generating the fusion proteins SAE1a:ECFP and SAE1b:ECFP, respectively. In these experiments, free EYFP and ECFP were used as controls (Figure 5B).

When SAE2:EYFP was co-expressed with SAE1a:ECFP, the fluorescence signal from EYFP was only detected in the nucleus, indicating that the E1 large subunit SAE2 localizes to the nucleus. On the contrary, also in the presence of SAE1a:ECFP, SAE2ΔCt:EYFP localized to the cytoplasm and the nuclear envelope, indicating that the SAE2 C-terminal region is required for SAE2 nuclear targeting. Identical results were obtained when SAE1b:ECFP was used instead of SAE1a:ECFP in co-expression experiments, suggesting that SAE2 subcellular localization is not affected by the small subunit SAE1 isoform (Figure 5C, left column). The same pattern was observed in all analyzed cells.

To further evaluate whether the SAE2 region containing the predicted nuclear localization signals at positions 579 and 582 was sufficient to determine nuclear localization, the SAE2 domain S426–E625 was fused to EYFP C-terminus

generating the construct EYFP:UFDCt. In onion epidermal cells, EYFP:UFDCt localized exclusively in the nucleus (Figure 5E). Consistently with a function of the SAE2 C-terminal tail in nuclear targeting, the fusion protein in which the C-terminal tail was removed, EYFP:UFD, displayed the same localization as the control EYFP (Figure 5F and 5B). In the last case, as in the control, signal was also detected in the nucleus since EYFP:UFD molecular weight, 40.5 kDa, is below the nuclear size exclusion limit.

Since the active E1 enzyme consists of the two subunits SAE2 and SAE1, it would be expected that both subunits display the same subcellular localization. When co-expressed with the nuclear-localized SAE2:EYFP, fluorescence from SAE1b:ECFP was observed exclusively in the nucleus in the 86% of analyzed cells. In addition to the nucleus, fluorescence from SAE1b:ECFP could also be detected in the cytoplasm in the other 14% of the cells. In contrast, in the presence of the nuclear-localized SAE2:EYFP, fluorescence from SAE1a:ECFP was detected in the nucleus and the cytoplasm in 83% of the analyzed cells. Exclusive nuclear localization of SAE1a:ECFP was restricted to the other 17% of the cells (Figure 5C, middle, and 5D). We excluded that these distinct distributions were the result of differential expression levels since SAE1a:ECFP and SAE1b:ECFP displayed comparable expression levels and localization, nuclear and cytoplasmic in all the analyzed cells, when co-expressed with SAE2ΔCt:EYFP (Figure 5C, middle, and 5D).

SAE1a Depletion Results in SUMO Conjugation Defects in Response to Stress

Since both isoforms of the SUMO-activating enzyme small subunit confer distinct conjugation efficiency *in vitro*, we explored the effect of SAE1a and SAE1b deletion in SUMO conjugation during stress responses. As *SAE1b* is duplicated in tandem in the genome, it is virtually impossible to obtain null T-DNA insertion *atsae1b* mutant plants. Thus, we focus on the analysis of the effect of SAE1a deletion in SUMO conjugation *in vivo*. We identified an *atsae1a* mutant line with a T-DNA insertion in the ninth exon of the *SAE1a* genomic sequence (Figure 6A). The T-DNA insertion site was corroborated by PCR combining primers complementary to *SAE1a* and T-DNA

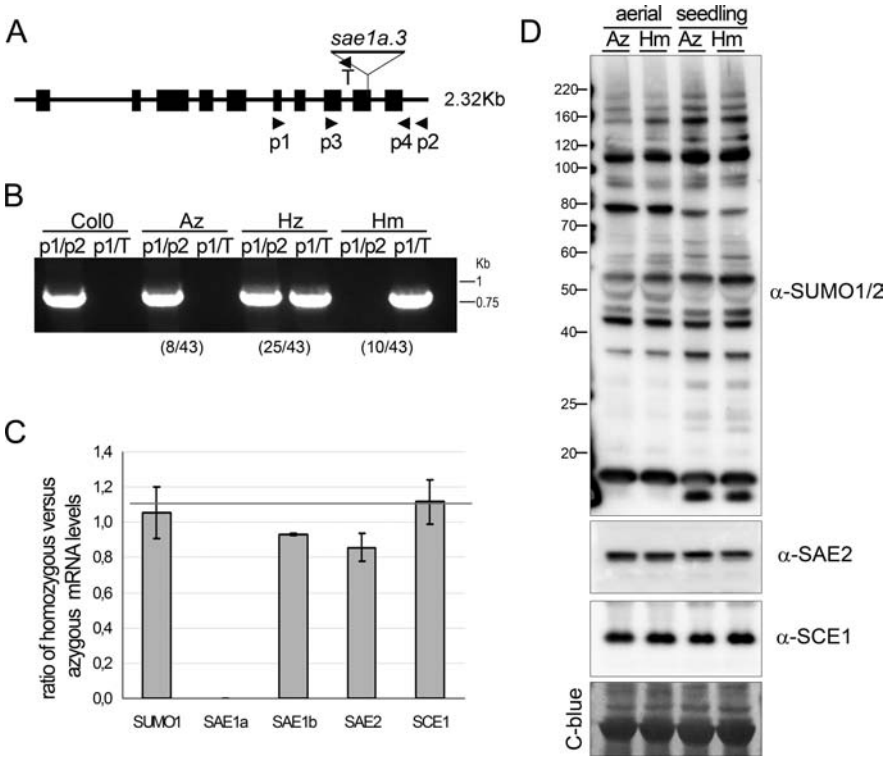


Figure 6. SUMO Conjugation in *atsae1a* Mutant Plants.
(A) *SAE1a* genomic region showing T-DNA insertion in the ninth exon. Annealing regions of primers used in (B) are represented by arrowheads.
(B) PCR products from plant genotype analysis were resolved in DNA agarose gel. Primers annealing to SAE1a genomic region (p1 and p2) and the T-DNA region (T) were used to identify azygous, heterozygous, and homozygous plants. Segregation results with respect to 43 individuals are indicated below the DNA agarose gel.
(C) mRNA levels corresponding to SUMO1, E1-activating enzyme (SAE1a, SAE1b, and SAE2), and E2-conjugating enzyme (SCE1) were determined in *atsae1a* azygous and homozygous lines. Collected data were normalized by using PR65 as a reference gene. Average values and standard deviation bars correspond to three biological replicates. Results are expressed as a ratio of the expression levels in homozygous plants versus azygous plants.
(D) Relative protein levels of free SUMO, SUMO conjugates, the E1-activating enzyme large subunit SAE2, and the E2-conjugating enzyme SCE1 in azygous and homozygous *sae1a* plants. 18 μg of total protein extracts from aerial parts and seedlings were resolved by SDS–PAGE and examined by immunoblot analysis with specific antibodies. A portion of the Coomassie blue (C-Blue)-stained membrane is shown as a loading control.

genomic sequences. The segregation analysis showed that *atsae1a* mutant was viable in homozygosis (Figure 6B). From the original segregating population, we selected azygous (#3.7) and homozygous lines (#3.9) for further characterization. Quantification of mRNA levels of other members of the sumoylation machinery showed that there were no major differences between the homozygous *atsae1a* and the corresponding azygous lines, excluding the presence of a compensatory mechanism affecting mRNA levels (Figure 6C). Also, no major differences could be observed among SUMO conjugates accumulation and SAE2 and SCE1 levels in *atsae1a* plants (Figure 6D). Under standard growth conditions, no significant developmental defects were observed in *atsae1a* plants (Supplemental Figure 5).

It is well established that heat and drought stresses induce a massive accumulation of high-molecular-weight SUMO conjugates (Kurepa et al., 2003), which presumably requires a highly active sumoylation machinery. Upon heat and drought stresses, *atsae1a* homozygous plants accumulated SUMO conjugates to a lower extent than azygous plants, which is consistent with the presence of defects in SUMO conjugation machinery (Figure 7A and 7B).

DISCUSSION

The E1-activating enzyme catalyzes the first step in SUMO conjugation and it has been proposed to have a role in SUMO isoform selection (Castaño-Miquel et al., 2011). In *Arabidopsis*, the E1-activating enzyme is present as two isoforms that differ in small subunit composition. The diversification of the small subunit into two conserved isoforms, SAE1a and SAE1b, is intriguing, since most of the E1 functional domains are located in the large subunit, SAE2. Structural analyses allowed the mapping of all functional domains in SAE2 as described for its human ortholog, although their conservation degree is not equally distributed among them. The most conserved is the adenylation domain, which is the only domain that depends on heterodimer formation. On the other hand, the cysteine and UFD domains are more divergent. It is remarkable that both the cysteine and the UFD domains establish non-covalent interactions with their cognate E2 conjugating enzyme, SCE1 (Wang et al., 2007, 2009, 2010), suggesting that sequence divergence in these regions between paralogs could be the result of protein–protein interaction surface optimization in each species. Other SAE2 regions containing regulatory elements display also a high divergence degree between paralogs, such as the Cys domain loop containing SUMO attachment sites and the C-terminal tail containing nuclear localization signals. These specific SAE2 regulatory region divergences suggest that the molecular mechanisms controlling SUMO conjugation *in vivo* might have diverged during evolution. Whether these differences are related to the distinct biological roles that SUMO plays in different organisms remains to be elucidated.

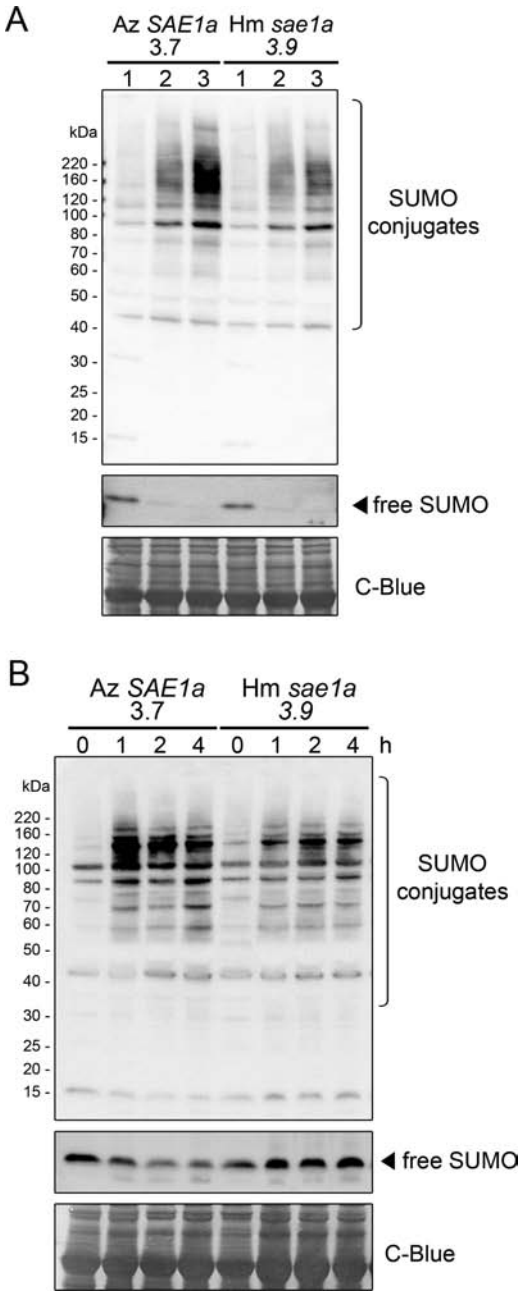


Figure 7. *atsae1a* Plants Display Defects in SUMO Conjugate Accumulation upon Abiotic Stress.

(A) SUMO conjugation upon heat-shock treatment. Azygous *SAE1a* (line #3.7) and homozygous *atsae1a* (lines #3.9) plants grown in liquid culture for 10 d were exposed to 42°C for 30 min (lane 2) and then returned at 22°C for 30 min for recovery (lane 3). Control plants were kept at 22°C (lane 1).

(B) SUMO conjugation upon dehydration treatment. Rosettes from 2-week-old plants grown in MS were detached from their roots and placed in a laminar flow hood. Samples were collected at the specified times. Total protein extracts were resolved by SDS–PAGE and SUMO conjugate levels examined by immunoblot analysis with anti-SUMO1 antibodies.

The E1 small subunit complexity is even higher at the gene organization level. *Arabidopsis thaliana* displays three *SAE1* genes (*SAE1a*, *SAE1b1*, and *SAE1b2*) resulting from two

Table 1. *Arabidopsis* SUMO Conjugation Machinery Components Analyzed in this Study.

Component	Name	TAIR accession	Protein MW
SUMO isoform 1	SUMO1	At4g26840	10.9 kDa
SUMO-activating enzyme subunit 2	SAE2	At2g21470	69.7 kDa
SUMO-activating enzyme subunit 1a	SAE1a	At4g24940	36.1 kDa
SUMO-activating enzyme subunit 1b	SAE1b	At5g50580 At5g50680	35.6 kDa
SUMO-activating enzyme	SCE1	At3g57870	18 kDa

duplications events, one preceding the emergence of the *Brassicacea* as indicated by the presence of two *SAE1* genes (*SAE1a* and *SAE1b*) in all family representatives, with the exception of *B. rapa*, and a more recent tandem duplication, which would have generated *SAE1b1* and *SAE1b2* genes. The fact that *SAE1b1* (At5g50580) and *SAE1b2* (At5g50680) codify for two identical proteins brought speculation about a possible annotation error (Novatchkova et al., 2012). Our results demonstrated the existence of this tandem duplication that affects 13 genes codifying for functionally unrelated proteins. Surprisingly, this duplication is not present in other *Arabidopsis thaliana* ecotypes analyzed, suggesting that tandem duplication giving raise to At5g50580 and At5g50680 occurred recently in the Col-0 lineage, after divergence from the closely related accessions/ecotypes Ler-1, C-24, Bur-0, and Kro-0. In addition to the evolutionary questions raised by this finding, it also opens the possibility to address the *in vivo* functional divergence of SAE1a/b isoforms in other *Arabidopsis thaliana* ecotypes containing a single gene codifying for each isoform.

The analysis of SAE1a and SAE1b primary sequences did not reveal the presence of any divergent domain that could suggest a functional specialization. Residues relevant to adenylation and those that undergo conformational changes during SUMO–SAE2 thioester bond formation are conserved in both isoforms. Nonetheless, SAE1a conferred higher conjugation efficiency relative to SAE1b under several incubation temperatures. In the case of SAE2 thioester bond formation, differences were restricted to reactions incubated at high temperatures, suggesting that SAE1 subunit composition would have a larger effect in events occurring downstream of SUMO–E1 thioester bond formation. SUMO activation and transfer to E2 involve dramatic conformational changes consisting of SAE2 rotations and unfolding and folding of specific regions in SAE2/SAE1 (Schulman, 2011). Considering that conformational flexibility is critical to biological activity (Freire, 2001), it is tempting to speculate that divergences in atomic interactions defining SAE1a and SAE1b three-dimensional structure might confer different stability properties to the E1 activating enzyme, which would affect SUMO conjugation.

The functional E1 nuclear localization signals have been analyzed using two approaches. In the first one, the full E1 heterodimer containing the native SAE2 or the SAE2 truncated form in which the C-terminal tail was removed were used. In the second, individual expression of the SAE2 C-terminal region not containing the predicted NLS located in the cysteine domain and including or not the C-terminal tail were analyzed. In both cases, the results indicated the crucial role of the SAE2 C-terminal tail targeting SAE2 to the nucleus, which is consistent with the nuclear localization of most SUMO conjugates (Saracco et al., 2007). Similarly, recent studies have shown that a nuclear localization signal located in the mammalian E1 large subunit is crucial for E1 nuclear targeting (Moutty et al., 2011). Even more recent is the finding that sumoylation of human SAE2 C-terminal tail is involved in SAE2 nuclear targeting (Truong et al., 2012b), suggesting that SAE2 nuclear localization is a highly regulated process involving different molecular strategies. All together, the emerging data support that the essential nuclear role of SUMO has been conserved through evolution. In contrast, the E1 small subunit SAE1 could be detected in the nucleus and the cytoplasm. This distribution was particularly evident for the isoform SAE1a, suggesting that a cytosolic reservoir of the SAE1 apoenzyme could exist. Previous studies showed that SAE2 transcript levels are significantly lower relative to SAE1a and SAE1b and it was proposed that a posttranslational mechanism would adjust the stoichiometry of E1 heterodimer subunit levels (Saracco et al., 2007). On the contrary, our results suggest that SAE1, and predominantly the SAE1a isoform, could exist as a free form. In this case and under SAE2-limiting concentrations, as suggested by transcriptional analysis, the most abundant SAE1 isoform would be preferentially assembled into the E1 holoenzyme, which could constitute a mechanism to regulate the SUMO conjugation rate *in vivo*.

In order to explore the relative contributions of SAE1a and SAE1b to SUMO conjugation *in vivo*, we searched for null *sae1a* and *sae1b* mutant plants. Under standard growth chamber conditions, the deletion of SAE1a did not result in obvious developmental alterations, consistently with a normal basal SUMO conjugate pattern displayed by *sae1a* homozygous plants relative to azygous plants. On the contrary, when sumoylation was analyzed in plants under heat and drought stresses, which induce a massive accumulation of SUMO conjugates, *sae1a* homozygous plants showed a reduction in SUMO conjugates accumulation. Overall, these results suggest that SAE1a is required for a complete functional sumoylation system under stress. Unfortunately, equivalent analysis could not be performed with SAE1b-null plants since *SAE1b* is duplicated in tandem in the *A. thaliana* Col0 genome and *sae1b* mutant plants cannot be isolated. As duplications within the SAE1 family are not restricted to *A. thaliana*, one hypothesis could be that duplications of the sumoylation system could constitute an advantage for its functionality.

There is increasing evidence supporting a regulatory role of the SUMO E1-activating enzyme in SUMO conjugation. In

mammals, low ROS levels inhibit SUMO conjugation by inducing a cross-linking between E1 and E2 enzymes active site (Bossis and Melchior, 2006). Also, the E1 large subunit SAE2 is modified by SUMO in two sites: one provides an inactive E1 pool that becomes activated by desumoylation under heat shock stress and the other contributes to SAE2 nuclear localization (Truong et al., 2012a, 2012b). In *Arabidopsis*, SUMO E1 was shown to participate in SUMO paralog selection (Castaño-Miquel et al., 2011). Overall, our results contribute to the notion that SUMO activation could constitute a regulatory step in the SUMO conjugation pathway. In addition to the conserved nuclear localization of the E1-activating enzyme, we have shown that the SAE1 isoform contained in the E1 holoenzyme has an effect in SUMO conjugation efficiency through a mechanism occurring most probably downstream of SUMO activation. *In vivo*, although we cannot discriminate between a gene dosage effect or the fact that SAE1a isoform confers a higher conjugation efficiency or both, our results show that SAE1a is necessary to maintain *in vivo* sumoylation homeostasis, pointing to a rate-limiting role of SUMO activation. E1 diversification is not restricted to *Arabidopsis* and it remains to be elucidated whether other regulatory mechanisms, such as SAE1 posttranslational modifications, may contribute to bringing functional diversity to other systems having a single E1 isoform.

METHODS

Plant Material and Growth Conditions

The *A. thaliana* T-DNA insertion mutant *atsae1a.3* in the Col-0 ecotype (Salk_060834) was identified at the TAIR database and obtained from the Arabidopsis Biological Resource Center (ABRC) (Alonso et al., 2003) as segregating T3 seeds. Homozygous *atsae1a.3* plants were identified by genomic PCR using the primers P1 and P2 for SAE1a genomic DNA amplification and P1 and LBb1 for T-DNA insertion detection. For *in vitro* cultures, seeds were stratified for 3 d, plated on Murashige and Skoog salts (Murashige and Skoog, 1962) (Duchefa), pH 5.7, supplemented with 0.8% BactoAgar (Difco), and transferred to a tissue culture room in a LD photoperiod (16 h light/8 h dark) at 22°C. For dehydration stress treatments, 16-day-old plants were cut near the stem–root junction and detached rosettes placed in a flow laminar hood for 1, 2, and 4 h. After treatments, plants were immediately frozen in liquid N_2 and stored at –80°C. For heat shock stress treatments, plants were grown in liquid medium Gamborg B5 (B5 vitamins and salts Duchefa, pH 5.7, glucose 20 g L^{-1}, MES 0.5 g L^{-1}) for 10 d at constant agitation (120 rpm). Plants were transferred to a water bath at 42°C for 30 min and then transferred to 22°C growth chamber for an additional 30 min of recovery. Plant samples were collected and frozen in liquid N_2. Plant tissues were collected from plants grown in soil except for young rosette and root tissues that were collected from 11-day-old seedlings grown on half-strength Murashige and Skoog salts supplemented with 1.5% BactoAgar.

Protein Extraction and Immunoblot

Anti-SUMO1/2, anti-SAE2, and anti-SCE1 antisera were generated in rabbits using full-length recombinant proteins (Cocalico). Plant tissue was ground in liquid nitrogen and proteins extracted with 50 mM Tris-HCl, pH 8, 150 mM NaCl, 0.2% Triton X-100, 1 mM PMSF, 1 μg ml^{-1} pepstatin, 1 μg ml^{-1} leupeptin, 2 mM *N*-ethylmaleimide, 10 mM iodoacetamide, 5 mM EDTA. 18 μg of total protein were resolved under reducing conditions by using SDS polyacrylamide gels and NuPage Novex 4–12% Bis/Tris Gels (Invitrogen). Proteins were transferred onto polyvinylidene difluoride (PVDF) membranes (Millipore), incubated with primary antibody overnight and secondary antibody, peroxidase-conjugated anti-rabbit (GE Healthcare), for 1 h at room temperature in TBST (20 mM Tris-HCl, pH 7.6, 20 mM NaCl, 0.1% (v/v) Tween20) supplemented with 3% non-fat dry milk. Peroxidase activity was developed in ECL Plus reagent (GE Healthcare) and chemiluminescence signal captured with the LAS-3000 imaging system (Fujifilm).

RNA Extraction and Quantitative Real-Time RT–PCR

Total RNA was extracted using the RNeasy Plant Mini kit (Qiagen) according to the manufacturer's recommendations. In the case of siliques, RNA was first extracted with TRIzol reagent (Invitrogen) and then cleaned up with a RNeasy spin column (Qiagen). Total RNA was treated with RNase-free DNase I (Promega) prior to retrotranscription. 1 μg of purified total RNA was retrotranscribed using oligodT as a primer with a Transcriptor High Fidelity cDNA Synthesis kit (Roche). Real-Time qPCR was carried out using Lightcycler® 480 SYBR Green I Master (Roche) in a Lightcycler® 480 (Roche) detection system following manufacturer's recommended amplification conditions. Reaction products were confirmed by melting curve analysis and by DNA agarose gel electrophoresis. PR65 (At1g13320) gene was chosen as a reference gene (Czechowski et al., 2005) to normalize the data from the different samples.

Transient Expression of Fluorescent Protein Fusions in Onion Cells

SAE2, *SAE2ΔCt SAE1a*, and *SAE1b* were fused in frame to the 5′ end of the coding sequences of yellow fluorescent protein (YFP) or cyan fluorescent protein (CFP) downstream of the 35S constitutive promoter. Onion epidermal cells were bombarded with 5 μg of each DNA construct using a helium biolistic gun (BIO-RAD). Treated epidermal cells were kept in the dark at room temperature for 16 h before analysis by confocal microscopy (Confocal Olympus FV 1000). YFP was excited with a 515-nm argon laser and images collected with a 550–630-nm range. CFP was excited with a 405-nm argon laser and images collected in the 460–500-nm range. Imaging of YFP and CFP imaging and transmissible light images collection were performed sequentially. Samples were scanned with the Z-stack mode and image stacks projection was calculated with ImageJ software (Rasband, 1997–2009).

In Vitro SUMO Conjugation and E1-Thioester Assays

Recombinant proteins were purified as previously described (Castaño-Miquel et al., 2011). In conjugation assays, we used the C-terminal tail of the *Arabidopsis* catalase 3 (419–472) fused to GST, GST:AtCAT3Ct. Reactions were carried out at the indicated temperatures in 25-µl reaction mixtures containing 1 mM ATP, 50 mM NaCl, 20 mM Hepes, pH 7.5, 0.1% Tween 20, 5 mM $MgCl_2$, 0.1 mM DTT, 2 µM SUMO, 0.5 µM AtSAE2/AtSAE1a, 0.5 µM AtSCE1, and 5 µM GST–AtCAT3Ct. After the specified incubation time, reactions were stopped by the addition of protein-loading buffer, incubated at 70°C for 10 min, and 10 µl aliquots were resolved by SDS–PAGE. Reaction products were detected by immunoblot analysis with anti-GST polyclonal antibodies (SIGMA, G7781). E1-thioester assays were performed at the specified temperatures in 25 µl reaction mixtures containing 1 mM ATP, 50 mM NaCl, 20 mM Hepes, pH 7.5, 0.1% Tween 20, 5 mM $MgCl_2$, 0.1 mM DTT, 2 µM SUMO, and 1 µM E1a or E1b. After 2 min, 15 µl aliquots were removed and analyzed by SDS–PAGE followed by Coomasie Fluor Orange staining according to the manufacturer's indications (Molecular Probes C-33250). As a thioester bond formation control, an aliquot of each reaction was treated with 100 mM DTT previously to loading into polyacrylamide gel. For SUMO conjugation efficiency quantifications, time-course reactions incubated at the same temperature and containing SAE1a or SAE1b isoforms were resolved in the same protein gel. After blotting to a PVDF (Millipore) membrane and incubated with anti-GST polyclonal antibodies (Sigma, G7781), luminescence signal generated by ECL Prime assay (GE Healthcare) was captured with a CCD camera (LAS4000, Fujifilm) and quantified with Multigauge software (Fujifilm). Each data point was normalized to the average of all data points obtained from each analyzed membrane in order to remove variability resulting from antibodies incubations and time exposure differences. The normalized values were used to calculate the corresponding slopes (relative luminescence signal versus time). The average slope from at least three independent experiments is shown in Figure 4. Relative SAE2~SUMO thioester levels indicate the amount of SAE2 in complex with SUMO relative to total SAE2 present in each data point. The average of three independent experiments is shown in Figure 4.

Structure Modeling, Sequence Identification, and Phylogenetic Analysis

Protein structure models were generated by using the SWISS-MODEL workspace (Arnold et al., 2006) on automated mode and specific structure template selection. SAE1a/b and SAE2 structures were predicted by using human E1 1Y8Q (2.25 Å) as a template. As a result of its low conservation degree, SAE2 Cys domain had to be modeled using the high-resolution human SAE2 Cys domain structure 2PX9 (1 Å). Models were assembled and images generated using PyMOL (DeLano, 2002).

Search for SAE1 and SAE2 sequences was performed throughout the whole genome of representative plant species using different BLAST-based programs (Altschul et al., 1997) in selected databases, including PLAZA 2.5 (http://bioinformatics.psb.ugent.be/plaza/) (Proost et al., 2009), phytozome v8.0 (www.phytozome.net/) (Goodstein et al., 2012), Cyanidioschyzon merolae Genome Project (http://merolae.biol.s.u-tokyo.ac.jp/), and NCBI (www.ncbi.nlm.nih.gov/). All hits were combined and redundant, unfinished, or non-full-length sequences were discarded. To confirm our data set of sequences as SAE1 and SAE2 sequences, we examined their fit to SAE1 and SAE2 key functional regions (Lois and Lima, 2005). Exon/intron location, distribution, and phases at the genomic sequences encoding for SAE1 and SAE2 sequences were examined through comparisons with the predicted encoded protein using GENEWISE (Birney et al., 2004).

Phylogenetic analyses were performed on the basis of multiple alignments of amino acid sequences obtained using MUSCLE (Edgar, 2004) and alignments edited with GeneDoc software (Nicholas and Nicholas, 1997). For SAE1 phylogenetic reconstruction, maximum-likelihood and Bayesian analyses were carried out using the JTT protein evolution model (Jones et al., 1992), heterogeneity of amino acid substitution rates corrected using a γ-distribution (G) with eight categories plus the proportion of invariant sites estimated by the data (I), selected by ProtTest v2.4 as the best-fitting amino acid substitution model according to the Akaike information criterion (Abascal et al., 2005). Similarly, maximum-likelihood and Bayesian analyses were run on SAE2 sequences using the JTT model plus G with eight categories. Bayesian analysis was implemented in MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001). Searches were run with four Markov chains for one million generations and sampling every 100th tree. After the stationary phase was reached (determined by the average standard deviation of split sequences approaching 0, which reflects the fact that independent tree samples became increasingly similar), the first 2500 trees were discarded as burn-in and a consensus tree was then constructed to evaluate Bayesian posterior probabilities on clades. Maximum-likelihood trees were constructed using PhyML v3.0 (Guindon and Gascuel, 2003; Guindon et al., 2010). Tree topology searching was optimized using the subtree pruning and regrafting option. The statistical support of the retrieved topology was assessed using the Shimodaira-Hasegawa-like approximate likelihood ratio test (Anisimova and Gascuel, 2006). Neighbor-joining phylogenetic analyses were conducted in MEGA 5.0 (Tamura et al., 2007). The evolutionary distances for neighbor-joining phylogenetic reconstruction were computed using the Poisson correction method. To obtain statistical support on the resulting clades, a bootstrap analysis with 1000 replicates was performed. Resulting trees were represented and edited using FigTree v1.3.1.

Accession Numbers

Assigned accession numbers for the studied genes are as follows: At5g55160 (SUMO2), At2g21470 (SAE2), At4g24940 (SAE1a), At5g50580 (SAE1b), At3g57870 (SCE1).

SUPPLEMENTARY DATA

Supplementary Data are available at *Molecular Plant Online.*

FUNDING

This work was supported by the European Research Council (grant ERC-2007-StG-205927) and the Spanish Ministry of Education and Science (grant BIO2008-01495 and CONSOLIDER CSD 2007-00036). L.C.M., S.M., I.T., and F.A. were supported by research contracts trough the CRAG. J.S. was supported by a pre-doctoral fellowship from the Ministry of Education and Science (BES-2005-6843). We also thank the Departament d'Innovació, Universitats i Empresa from the Generalitat de Catalunya (Xarxa de Referència en Biotecnologia and 2009SGR 09626) for substantial support.

ACKNOWLEDGMENTS

The technical support from members of the Greenhouse, Microscopy, and Genomics facilities at CRAG is greatly appreciated. We greatly thank Cristina Cañadas for technical support at LML lab and Dr Javier Forment Millet for bioinformatics technical support at IBMCP. No conflict of interest declared.

REFERENCES

Abascal, F., Zardoya, R., and Posada, D. (2005). ProtTest: selection of best-fit models of protein evolution. Bioinformatics. **21**, 2104–2105.

Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., et al. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. Science. **301**, 653–657.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**, 3389–3402.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. **55**, 539–552.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. **22**, 195–201.

Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and Genomewise. Genome Res. **14**, 988–995.

Blanc, G., Hokamp, K., and Wolfe, K.H. (2003). A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Res. **13**, 137–144.

Bossis, G., and Melchior, F. (2006). Regulation of SUMOylation by reversible oxidation of SUMO conjugating enzymes. Mol. Cell. **21**, 349–357.

Castaño-Miquel, L., Seguí, J., and Lois, L.M. (2011). Distinctive properties of *Arabidopsis* SUMO paralogues support the *in vivo* predominant role of AtSUMO1/2 isoforms. Biochem. J. **436**, 581–590.

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

Chaikam, V., and Karlson, D.T. (2010). Response and transcriptional regulation of rice SUMOylation system during development and stress conditions. BMB Rep. **43**, 103–109.

Chen, C.-C., Chen, Y.-Y., Tang, I.C., Liang, H.-M., Lai, C.-C., Chiou, J.-M., and Yeh, K.-C. (2011). *Arabidopsis* SUMO E3 ligase SIZ1 is involved in excess copper tolerance. Plant Physiol.

Couvreur, T.L., Franzke, A., Al-Shehbaz, I.A., Bakker, F.T., Koch, M.A., and Mummenhoff, K. (2010). Molecular phylogenetics, temporal diversification, and principles of evolution in the mustard family (Brassicaceae). Mol. Biol. Evol. **27**, 55–71.

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M.K., and Scheible, W.R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in *Arabidopsis*. Plant Physiol. **139**, 5–17.

DeLano, W.L. (2002). The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA, USA, available online at www.pymol.org.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**, 1792–1797.

Freire, E. (2001). The thermodynamic linkage between protein structure, stability, and function. Methods Mol. Biol. **168**, 37–68.

Gareau, J.R., and Lima, C.D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat. Rev. Mol. Cell Biol. **11**, 861–871.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., et al. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. **40**, D1178–D1186.

Guindon, S., and Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**, 696–704.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. **59**, 307–321.

Hermkes, R., Fu, Y.F., Nurrenberg, K., Budhiraja, R., Schmelzer, E., Elrouby, N., Dohmen, R.J., Bachmair, A., and Coupland, G. (2011). Distinct roles for *Arabidopsis* SUMO protease ESD4 and its closest homolog ELS1. Planta. **233**, 63–73.

Huelsenbeck, J.P., and Ronquist, F. (2001). MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. **17**, 754–755.

Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992). The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. **8**, 275–282.

Kim, J.G., Taylor, K.W., Hotson, A., Keegan, M., Schmelz, E.A., and Mudgett, M.B. (2008). XopD SUMO protease affects host transcription, promotes pathogen growth, and delays symptom development in Xanthomonas-infected tomato leaves. Plant Cell. **20**, 1915–1929.

Kurepa, J., Walker, J.M., Smalle, J., Gosink, M.M., Davis, S.J., Durham, T.L., Sung, D.Y., and Vierstra, R.D. (2003). The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis: accumulation of SUMO1 and –2 conjugates is increased by stress. J. Biol. Chem. **278**, 6862–6872.

Lee, I., and Schindelin, H. (2008). Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. Cell. **134**, 268–278.

Lois, L.M. (2010). Diversity of the SUMOylation machinery in plants. Biochem. Soc. Trans. **38**, 60–64.

Lois, L.M., and Lima, C.D. (2005). Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. EMBO J. **24**, 439–451.

Long, Y., Zhao, L., Niu, B., Su, J., Wu, H., Chen, Y., Zhang, Q., Guo, J., Zhuang, C., Mei, M., et al. (2008). Hybrid male sterility in rice controlled by interaction between divergent alleles of two adjacent genes. Proc. Natl Acad. Sci. U S A. **105**, 18871–18876.

Miura, K., and Hasegawa, P.M. (2010). Sumoylation and other ubiquitin-like post-translational modifications in plants. Trends Cell Biol. **20**, 223–232.

Miura, K., Rus, A., Sharkhuu, A., Yokoi, S., Karthikeyan, A.S., Raghothama, K.G., Baek, D., Koo, Y.D., Jin, J.B., Bressan, R.A., et al. (2005). The *Arabidopsis* SUMO E3 ligase SIZ1 controls phosphate deficiency responses. Proc. Natl Acad. Sci. U S A. **102**, 7760–7765.

Moutty, M.C., Sakin, V., and Melchior, F. (2011). Importin alpha/beta mediates nuclear import of individual SUMO E1 subunits and of the holo-enzyme. Mol. Biol. Cell. **22**, 652–660.

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol. Plant. **15**, 473–497.

Murtas, G., Reeves, P.H., Fu, Y.F., Bancroft, I., Dean, C., and 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.

Nicholas, K.B., and Nicholas, H.B., Jr. (1997). GeneDoc: a tool for editing and annotating multiple sequence alignments. Available online at www.psc.edu/biomed/genedoc.

Novatchkova, M., Tomanov, K., Hofmann, K., Stuible, H.-P., and Bachmair, A. (2012). Update on sumoylation: defining core components of the plant SUMO conjugation system by phylogenetic comparison. New Phytol. **195**, 23–31.

Olsen, S.K., Capili, A.D., Lu, X., Tan, D.S., and Lima, C.D. (2010). Active site remodelling accompanies thioester bond formation in the SUMO E1. Nature. **463**, 906–912.

Park, B.S., Song, J.T., and Seo, H.S. (2011). *Arabidopsis* nitrate reductase activity is stimulated by the E3 SUMO ligase AtSIZ1. Nature Commun. **2**, 400.

Park, H.C., Kim, H.U.N., Koo, S.C., Park, H.J., Cheong, M.S., Hong, H., Baek, D., Chung, W.S., Kim, D.H., Bressan, R.A., et al. (2010). Functional characterization of the SIZ/PIAS-type SUMO E3 ligases, OsSIZ1 and OsSIZ2 in rice. Plant, Cell Environ. **33**, 1923–1934.

Proost, S., Van Bel, M., Sterck, L., Billiau, K., Van Parys, T., Van de Peer, Y., and Vandepoele, K. (2009). PLAZA: a comparative genomics resource to study gene and genome evolution in plants. Plant Cell. **21**, 3718–3731.

Rasband, W.S. (1997–2009). ImageJ. US National Institutes of Health, Bethesda, MD, USA, available online at http://rsb.info.nih.gov/ij/.

Saracco, S.A., Miller, M.J., Kurepa, J., and Vierstra, R.D. (2007). Genetic analysis of SUMOylation in *Arabidopsis*: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. Plant Physiol. **145**, 119–134.

Schulman, B.A. (2011). Twists and turns in ubiquitin-like protein conjugation cascades. Protein Sci. **20**, 1941–1954.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. Mol. Biol. Evol.

Thangasamy, S., Guo, C.L., Chuang, M.H., Lai, M.H., Chen, J.C., and Jauh, G.Y. (2011). Rice SIZ1, a SUMO E3 ligase, controls spikelet fertility through regulation of anther dehiscence. New Phytol. **189**, 869–882.

Truong, K., Lee, T., and Chen, Y. (2012a). SUMO modification of the E1 Cys domain inhibits its enzymatic activity. J. Biol. Chem.

Truong, K., Lee, T.D., Li, B., and Chen, Y. (2012b). Sumoylation of SAE2 C terminus regulates SAE nuclear localization. J. Biol. Chem. **287**, 42611–42619.

Wang, H.D., Makeen, K., Yan, Y., Cao, Y., Sun, S.B., and Xu, G.H. (2011). OsSIZ1 regulates the vegetative growth and reproductive development in rice. Plant Mol. Biol. Rep. **29**, 411–417.

Wang, J., Hu, W., Cai, S., Lee, B., Song, J., and Chen, Y. (2007). The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications. Mol. Cell. **27**, 228–237.

Wang, J., Lee, B., Cai, S., Fukui, L., Hu, W., and Chen, Y. (2009). Conformational transition associated with E1-E2 interaction in small ubiquitin-like modifications. J. Biol. Chem. **284**, 20340–20348.

Wang, J., Taherbhoy, A.M., Hunt, H.W., Seyedin, S.N., Miller, D.W., Miller, D.J., Huang, D.T., and Schulman, B.A. (2010). Crystal structure of UBA2ufd–Ubc9: insights into E1–E2 interactions in SUMO pathways. PloS One. **5**, e15805.

Wilkinson, K.A., and Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. Biochem. J. **428**, 133–145.

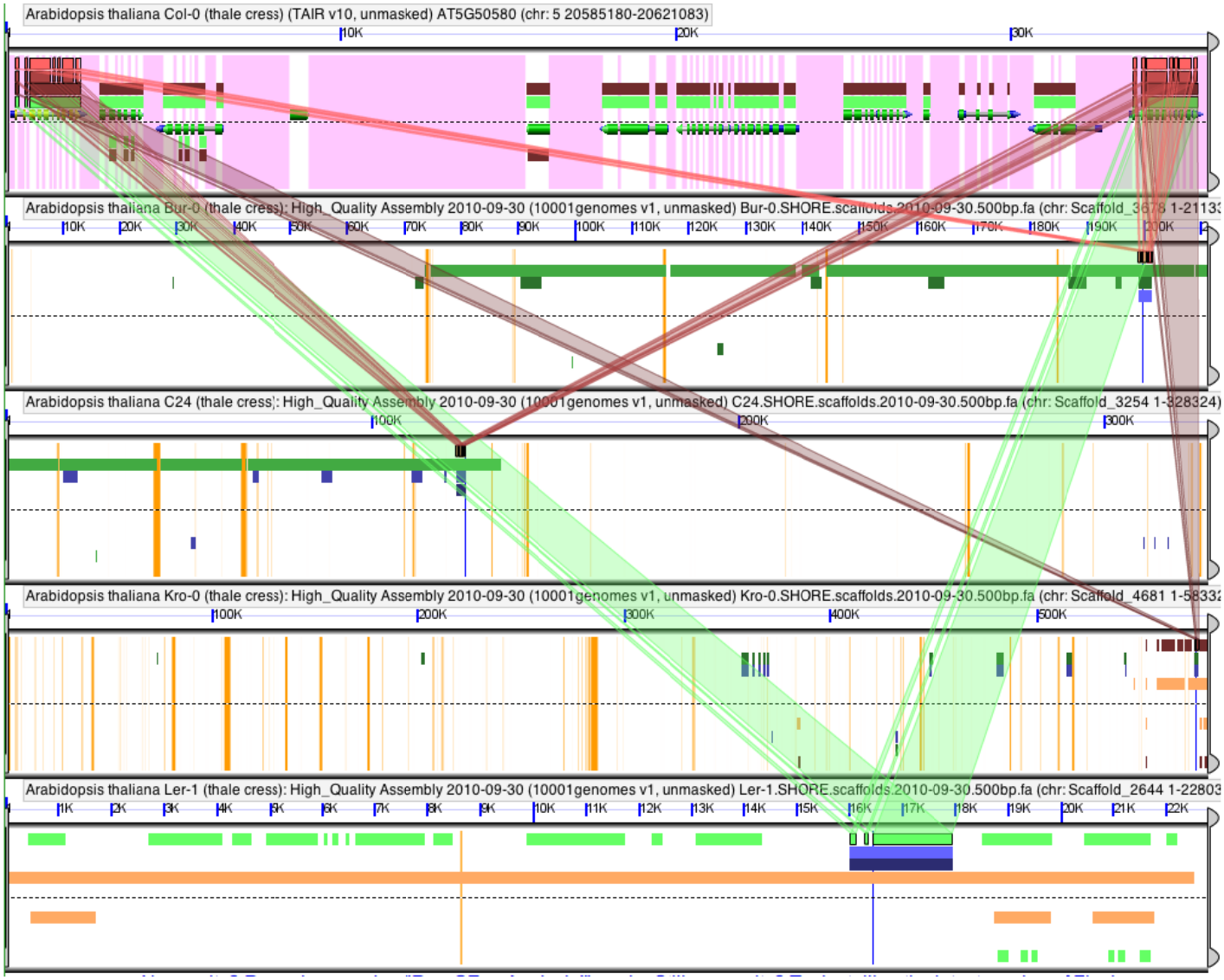


Figure S1. Syntenic map of *SAE1B* duplication.

We have used the nucleotide sequences of *Arabidopsis thaliana* Col-0 gene tandem duplicates AT5G50580 and AT5G50680 as queries in independent BLASTN searches (http://www.1001genomes.org/) of the whole sequenced genomes of *A. thaliana* accessions/ecotypes Ler-1, C24, Bur-0 and Kro-0, (Cao et al. 2011) from 1001 Genomes Project (http://www.1001genomes.org/) (Schneeberger et al. 2011). In all four genomes, the same scaffold region was retrieved as the best hit using both queries (Scaffold_2644, Ler-1; Scaffold_3254, C24; Scaffold_3678, Bur-0; Scaffold_4681, Kro-0). A comparative genomics analysis was also performed using the GEVO tool from the platform COGE

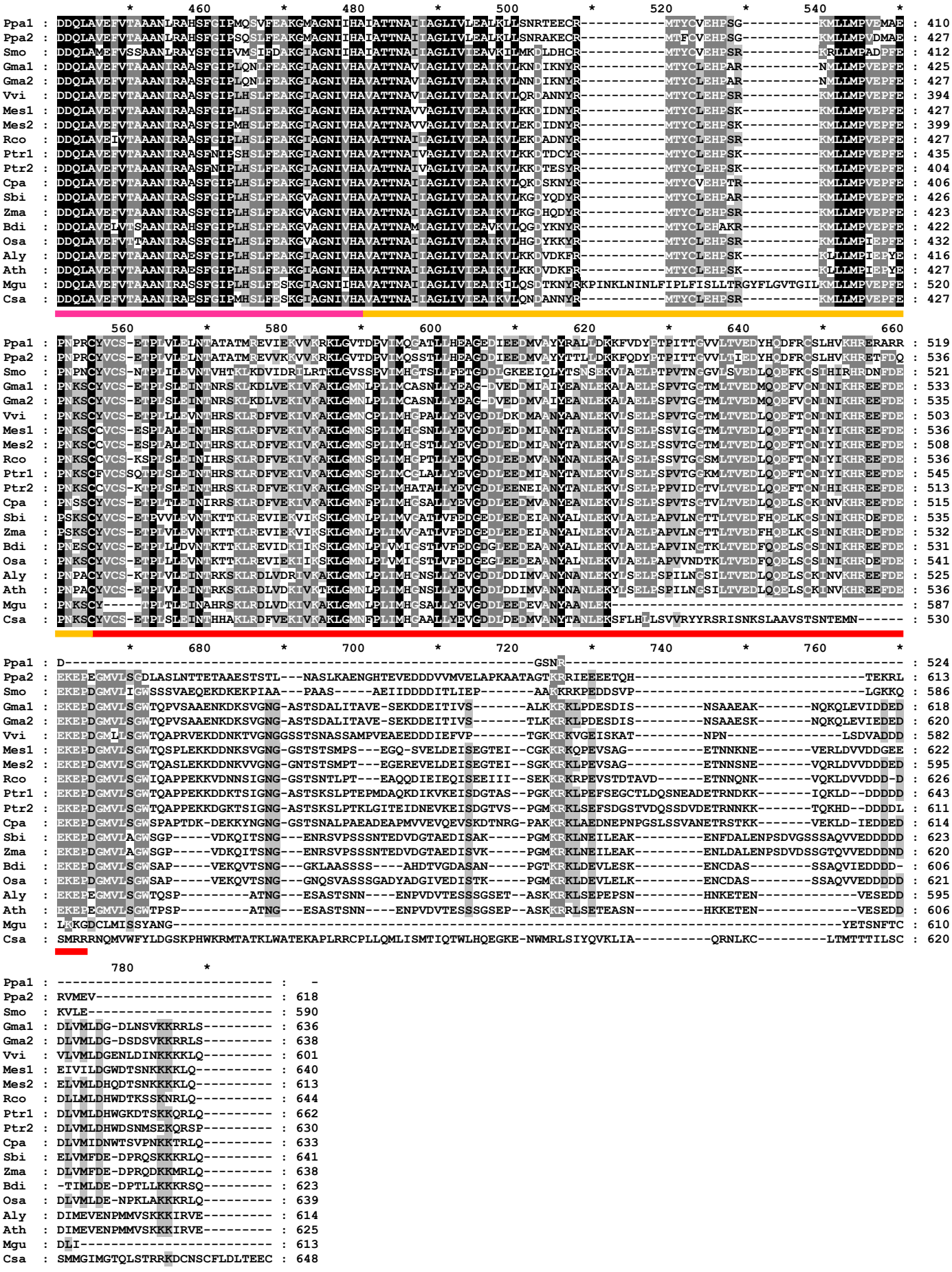
(http://genomevolution.org/CoGe/GEvo.pl). The genomic 33kb region of *A. thaliana* Col-0 extending from AT5G50580 to AT5G50680 was compared to the corresponding syntenic regions in *A. thaliana*Ler-1, C24, Bur-0 and Kro-0accessions/ecotypes. For the sake of clarity, non-cds regions were masked from *A. thaliana* Col-0 genomic region. The analysis may be regenerated following the link http://genomevolution.org/r/67hu. HSPs (High-Scoring Segment Pairs) between AT5G50580 and AT5G50680 coding sequences matched single/individual syntenic genomic regions in Ler-1, C24, Bur-0 and Kro-0 (showed by the connectors in the Figure). From these results it can be concluded that tandem duplication giving raise to AT5G50580 and AT5G50680 occurred recently in the Col-0 lineage, after divergence from the closely related accessions/ecotypes Ler-1, C-24, Bur-0 and Kro-0.We haven't found further information about evolutionary diversification and divergence times of these different accessions/ecotypes to give a more precise estimate of the tandem duplication timing.

Literature cited

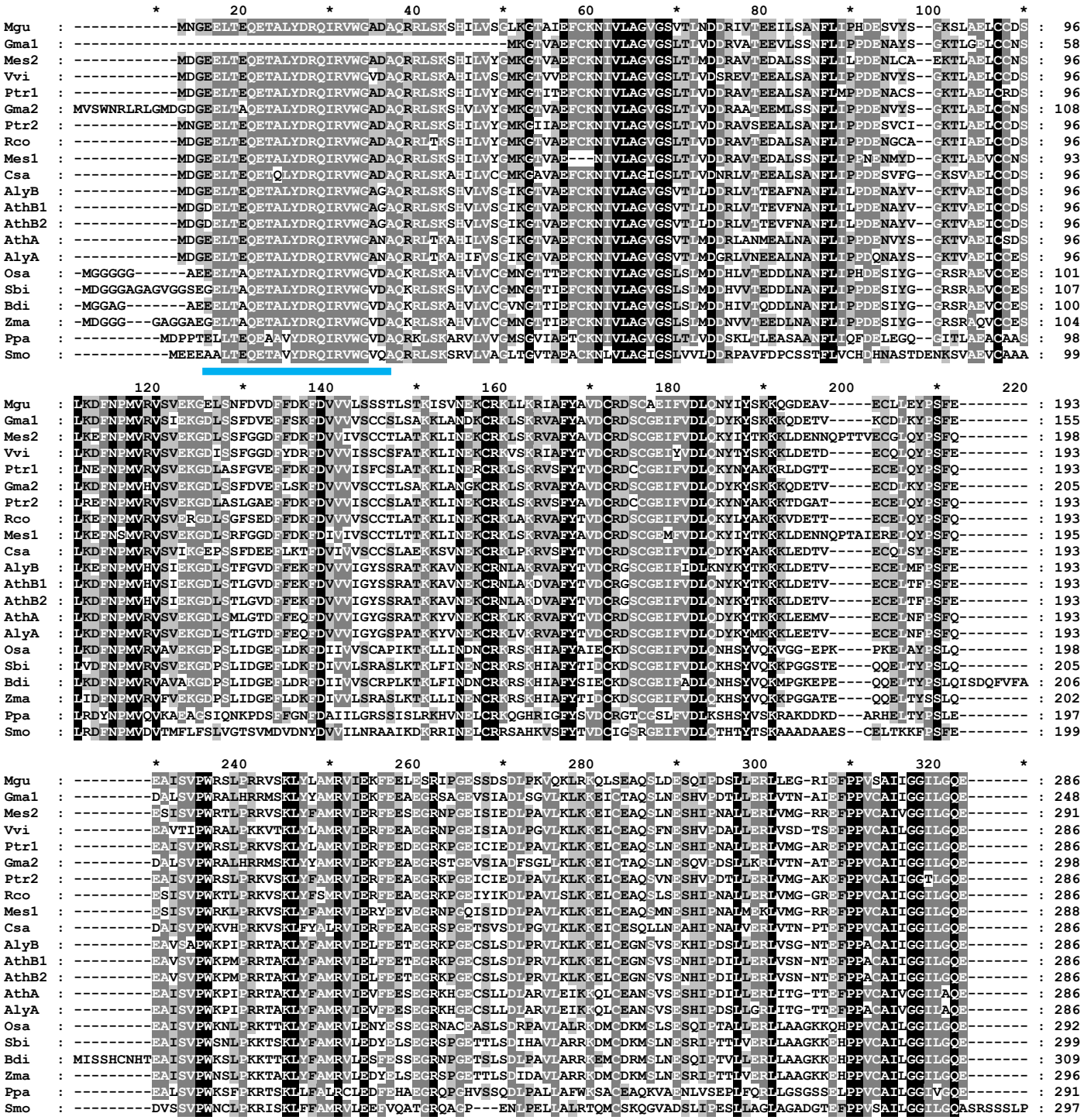
Cao, J, K Schneeberger, S Ossowski, et al. 2011. Whole-genome sequencing of multiple Arabidopsis thaliana populations. Nat Genet 43:956-963.

Schneeberger, K, S Ossowski, F Ott, et al. 2011. Reference-guided assembly of four diverse Arabidopsis thaliana genomes. Proc Natl Acad Sci U S A 108:10249-10254.


Ppa1
Ppa2
Smo
Gma1
Gma2
Vvi
Mes1
Mes2
Rco
Ptr1
Ptr2
Cpa
Sbi
Zma
Bdi
Osa
Aly
Ath
Mgu
Csa



B


Mgu
Gma1
Mes2
Vvi
Ptr1
Gma2
Ptr2
Rco
Mes1
Csa
AlyB
AthB1
AthB2
AthA
AlyA
Osa
Sbi
Bdi
Zma
Ppa
Smo

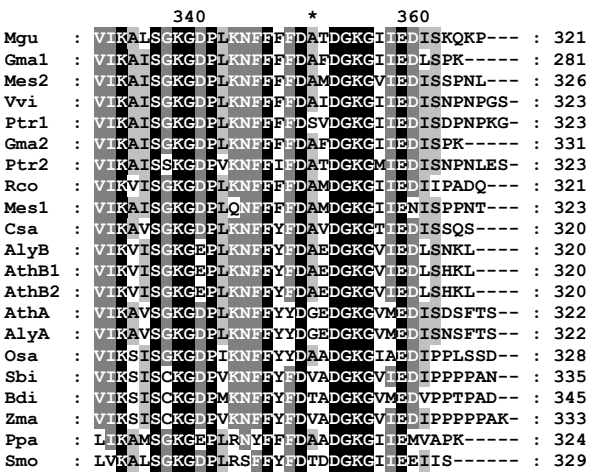

340 * 360
Mgu : VIKALSGKGDPLKNFFFFDATDGKGILEDISKQKP--- : 321
Gma1 : VIKAISGKGDPLKNFFFFDAFDGKGILEDLSPK----- : 281
Mes2 : VIKAISGKGDPLKNFFFFDAMDGKGVLEDISSPNL--- : 326
Vvi : VIKAISGKGDPLKNFFFFDAIDGKGILEDISNPNPGS- : 323
Ptr1 : VIKAISGKGDPLKNFFFFDSVDGKGILEDISDPNPKG- : 323
Gma2 : VIKAISGKGDPLKNFFFFDAFDGKGILEDISPK----- : 331
Ptr2 : VIKAISSKGDPVKNFFLFDATDGKGMLEDISNPNLES- : 323
Rco : VIKVISGKGDPLKNFFFFDAMDGKGILEDIIPADQ--- : 321
Mes1 : VIKAISGKGDPLQNFFFFDAMDGKGILENISPPNT--- : 323
Csa : VIKAVSGKGDPLKNFFYFDAVDGKGTLEDISSQS---- : 320
AlyB : VIKVISGKGEPLKNFFYFDAEDGKGVLEDLSNKL---- : 320
AthB1 : VIKVISGKGEPLKNFFYFDAEDGKGVLEDLSHKL---- : 320
AthB2 : VIKVISGKGEPLKNFFYFDAEDGKGVLEDLSHKL---- : 320
AthA : VIKAVSGKGDPLKNFFYYDGEDGKGVMEDISDSFTS-- : 322
AlyA : VIKAVSGKGDPLKNFFYYDGEDGKGVMEDISNSFTS-- : 322
Osa : VIKSISGKGDPIKNFFYYDAADGKGIAEDIPPLSSD-- : 328
Sbi : VIKSISCKGDPVKNFFYFDVADGKGVLEDIPPPPAN-- : 335
Bdi : VIKSISCKGDPMKNFFYFDTADGKGVMEDVPPTPAD-- : 345
Zma : VIKSISCKGDPVKNFFYFDVADGKGVLEDIPPPPPAK- : 333
Ppa : LLKAMSGKGEPLRNYFFFDAADGKGILEMVAPK----- : 324
Smo : LVKALSGKGDPLRSFFYFDTDDGKGILEEIIS------ : 329

Figure S2. SAE2 (A) and SAE1 (B) sequence alignments used for phylogenetic tree construction. Black background and white letters corresponds to 100% sequence identity. Gaps in the alignment due to insertions or deletions are indicated by dashed lines. Residue numbers are shown to the right of the sequences. Distribution of SAE2 functional domains is indicated by a line below the alignment colored as follows: adenylation domain in yellow, cysteine domain in pink and UFD domain in red. SAE1 region that undergoes conformational changes between the open and closed E1 conformations is indicated by a blue line below the alignment. Aly, *Arabidopsis lyrata*; Ath, *Arabidopsis thaliana*; Bdi, *Brachypodium distachyon*; Cpa, *Carica papaya*; Csa, *Cucumis sativus*; Gma, *Glycine max*; Mes, *Manihot esculenta*; Mgu, *Mimulus guttatus*; Mtr, *Medicago truncatula*; Osa, *Oryza sativa*; Ppa, *Physcomitrella patens*; Ptr, *Populus trichocarpa*; Rco, *Ricinus communis*; Sbi, *Sorghum bicolor*; Smo, *Selaginella moellendorffii*; Vvi, *Vitis vinifera*; Zma, *Zea mays*.

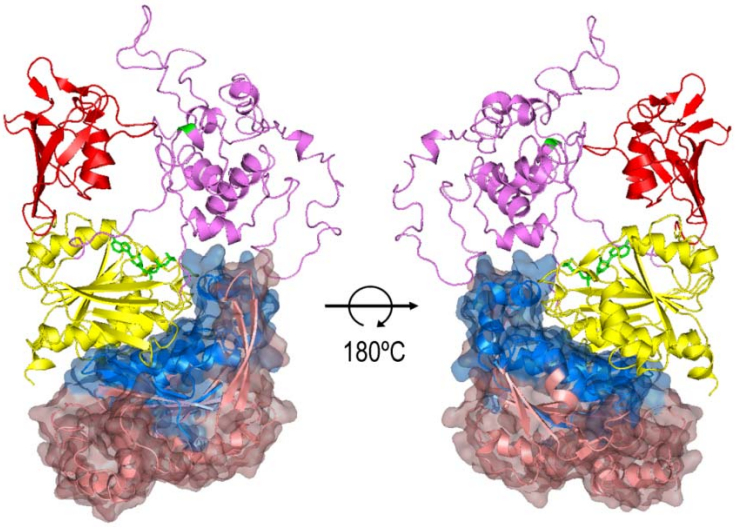


Figure S3. Distribution of low conserved regions between Arabidopsis and human SUMO activating enzyme E1. A, E1a (SAE2/SAE1a complex) predicted structure is shown as a ribbon diagram. SAE2 functional domains are represented in yellow (adenylation domain), violet (cysteine domain) and red (Ubiquitin fold domain, UFD). SAE1a surface is also shown. Blue indicates the most conserved regions (residues 6-106+271-317; 41% identity) and salmon, the most divergent regions (residues 107-270; 21% identity).

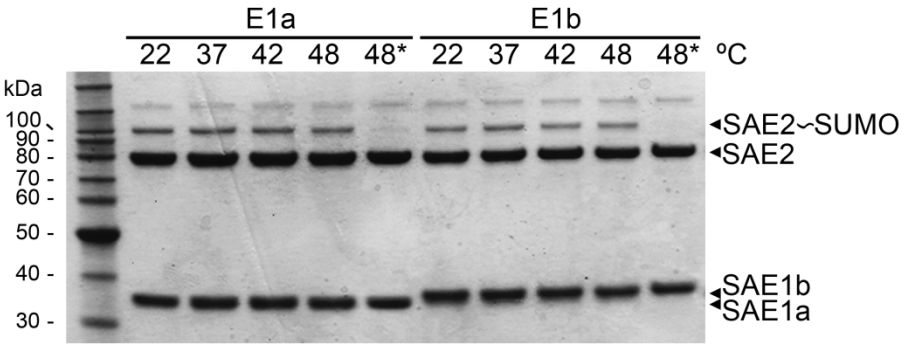


Figure S4. SAE2~SUMO thioester assay replicate. SAE2~SUMO complex is sensitive to DTT treatment (48* lane).

Figure S5. 5-week old azygous and *atsae1a* homozygous plants grown under long day conditions in growth chambers

Table S1.

List of genes contained in the analyzed tandem duplicated region of chromosome 5.

Locus identifier	Gene description
AT5G50530	CBS / octicosapeptide/Phox/Bemp1 (PB1) domains-containing protein
AT5G50640	
AT5G50600	Hydroxysteroid dehydrogenase HSD1
AT5G50700	
AT5G50580	SUMO activating enzyme 1b
AT5G50680	
AT5G50590	Hydroxysteroid dehydrogenase HSD4
AT5G50690	
AT5G50510	Molecular chaperone Hsp40/DnaJ family protein
AT5G50620	
AT5G50520	Major facilitator superfamily protein
AT5G50630	
AT5G50540	unknown protein
AT5G50645	
AT5G50550	Transducin/WD40 repeat-like superfamily protein
AT5G50650	
AT5G50560	Protein of unknown function (DUF 3339)
AT5G50660	
AT5G50562	This gene encodes a small protein and has either evidence of transcription or purifying selection (it is duplicated but not annotated in MFB16)
AT5G50565	unknown protein
AT5G50665	
AT5G50570	Squamosa promoter-binding protein-like (SBP domain) transcription factor family protein
AT5G50670	
AT5G50610	unknown protein
AT5G50710	

Table S2.

List of SAE2 homologous sequences used in the phylogenetic analysis. List is arranged according to the gene identifier alphabetical order.

Gene Identifier	Species	Lineage
AL4G00660	*Arabidopsis lyrata*	Dicots-Brassicaceae
AT2G21470	*Arabidopsis thaliana*	Dicots-Brassicaceae
BD1G23100	*Brachypodium distachyon*	Monocots
Bra030297	*Brassica rapa*	Dicots-Brassicaceae
Bra031191	*Brassica rapa*	Dicots-Brassicaceae
Carubv10022825m.g	*Capsella rubella*	Dicots-Brassicaceae
CME143C	*Cyanidioschyzon merolae*	Red Algae
CP01026G00010	*Carica papaya*	Dicots
Cucsa.109340	*Cucumis sativus*	Dicots
Eucgr.G01699	*Eucalyptus grandis*	Dicots
FV2G01100	*Fragaria vesca*	Dicots
GM12G36450	*Glycine max*	Dicots
GM13G27140	*Glycine max*	Dicots
MD00G069110	*Malus domestica*	Dicots
MD10G007620	*Malus domestica*	Dicots
ME02994G00460	*Manihot esculenta*	Dicots
ME02994G00470	*Manihot esculenta*	Dicots
MRCC299_13G04660	*Micromonas sp. RCC299*	Green Algae
orange1.1g006294m.g	*Citrus sinensis*	Dicots
OS07G39780	*Oryza sativa ssp. japonica*	Monocots
OSINDICA_07G32090	*Oryza sativa ssp. indica*	Monocots
Pp1s197_66V6	*Physcomitrella patens*	Basal Land Plants
Pp1s54_122V6	*Physcomitrella patens*	Basal Land Plants
ppa002704m.g	*Prunus persica*	Dicots
PT04G15680	*Populus trichocarpa*	Dicots
PT09G12110	*Populus trichocarpa*	Dicots
RC30131G02360	*Ricinus communis*	Dicots
SB02G037850	*Sorghum bicolor*	Monocots
Si029042m.g	*Setaria italica*	Monocots
SM00105G00100	*Selaginella moellendorffii*	Basal Vascular Plants
TC02G007630	*Theobroma cacao*	Dicots
Thhalv10000089m.g	*Thellungiella halophila*	Dicots-Brassicaceae
VC00016G01070	*Volvox carteri*	Green Algae

Table S3.

List of SAE1 homologous sequences used in the phylogenetic analysis. List is arranged according to the gene identifier alphabetical order.

Gene Identifier	Species	Lineage
294464849	*Picea sitchensis*	Gymnosperms
Aco_AcoGoldSmith_v1.008486m.g	*Aquilegia coerulea Goldsmith*	Dicots
AL7G17210	*Arabidopsis lyrata*	Dicots-Brassicaceae
AL8G14820	*Arabidopsis lyrata*	Dicots-Brassicaceae
AT4G24940	*Arabidopsis thaliana*	Dicots-Brassicaceae
AT5G50580	*Arabidopsis thaliana*	Dicots-Brassicaceae
AT5G50680	*Arabidopsis thaliana*	Dicots-Brassicaceae
BD4G18870	*Brachypodium distachyon*	Monocots
Bra013848	*Brassica rapa*	Dicots-Brassicaceae
Carubv10005251m.g	*Capsella rubella*	Dicots-Brassicaceae
Carubv10026795m.g	*Capsella rubella*	Dicots-Brassicaceae
Ccl_0.9_015767m.g	*Citrus clementina*	Dicots
CMC124C	*Cyanidioschyzon merolae*	Red Algae
Csa_392180.1	*Cucumis sativus*	Dicots
Csi_1.1g020575m.g	*Citrus sinensis*	Dicots
Egr_K02707.1	*Eucalyptus grandis*	Dicots
FV0G09730	*Fragaria vesca*	Dicots
GM05G38190	*Glycine max*	Dicots
GM08G01440	*Glycine max*	Dicots
ME01701G00210	*Manihot esculenta*	Dicots
ME09876G00390	*Manihot esculenta*	Dicots
Mgu_mgv1a010151m.g	*Mimulus guttatus*	Green Algae
MRCC299_03G00430	*Micromonas sp. RCC299*	Green Algae
OL19G00060	*Ostreococcus lucimarinus*	Green Algae
OS11G30410	*Oryza sativa ssp. japonica*	Monocots
OSINDICA_11G21650	*Oryza sativa ssp. indica*	Monocots
OT20G00870	*Ostreococcus tauri*	Green Algae
PP00027G01760	*Physcomitrella patens*	Basal Land Plants
Ppe_ppa009164m.g	*Prunus persica*	Dicots
PT12G09700	*Populus trichocarpa*	Dicots
PT15G10160	*Populus trichocarpa*	Dicots
RC30147G03790	*Ricinus communis*	Dicots
SB05G017710	*Sorghum bicolor*	Monocots
SM00046G01840	*Selaginella moellendorffii*	Basal Vascular Plants
TC03G015120	*Theobroma cacao*	Dicots
Thhalv10025730m.g	*Thellungiella halophila*	Dicots-Brassicaceae
VC00042G00710	*Volvox carteri*	Green Algae
VV17G08240	*Vitis vinifera*	Dicots
ZM04G04670	*Zea mays*	Monocots

Table S4

List of primers used in this study.

Name	Gene/Contig	Sequence 5'-3'
P1	SAE1a	AGAGATGGTGGAATGCGAAC
P2	SAE1a	GCCCACTGGTTGTTTAAGAG
P3	SAE1a	GAAGGGCGTAAACATGGAGA
P4	SAE1a	GGCCCACTGGTTGTTTAAGA
SM004	SAE2	TTCTTCTTCATCAAAGAATCCC
SM019	SAE2	TAAATGTCAAGCACAGGGAAG
SM020	SAE2	CTCCCATTCTTAACGGAAGC
SM021	SAE2	TTCCATTTCACTGAATCACG
FA715	MBA10/MFB16	ATGTCTAAGACTTCGACAAGG
FA716	MBA10/MFB16	AAGTTGTCTGAGTATCGTGG
FA717	MBA10	TTCTGGATCATTTTCACCTCC
FA718	MFB16	GACATCATCTTTGACCCTAG

Primers used in qPCR.

Gene	Direction	Sequence 5'-3'
SUMO1	Forward	TGTGTGTGATATGTGAATGGTTAAGG
	Reverse	TCTGCAAATCCTAATGCCGAT
SAE2	Forward	TCAGGATCCGAGCCTGCTTC
	Reverse	TACTGGCCTCGGTCTCAGAG
SAE1a	Forward	CCGCCAGTTTGTGCCATTG
	Reverse	GCTTTGATCACCTCCTGTGC
SCE1	Forward	GCTCGTGGTCGTTTAGCTGAA
	Reverse	GCCACAAAACCATGAGGATGA
SAE1b	Forward	CGAGGACCTATCCCACAAGC
	Reverse	GCGTTTCACAGCACCTTTCAT
PR65	Forward	TAACGTGGCCAAAATGATGC
	Reverse	GTTCTCCACAACCGCTTGGT

CAPÍTULO-III

SUMOylation inhibition *in vivo* is achieved through the expression of the SUMO E1 UFDCt domain

SUMOylation inhibition *in vivo* is achieved through the expression of the SUMO E1 UFDCt domain

Laura Castaño-Miquel*, Inês Teixeira*, Josep Seguí, Abraham Mas, Natalia Rodrigo, Gaelle La Verde and L. Maria Lois[1].

1 Development Program, CSIC- Centro de Investigaciones Científicas. Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.
***** These authors contributed equally to this work

Resumen. La modificación de proteínas por SUMO es un mecanismo fundamental que regula una amplia variedad de procesos celulares en eucariotas. El estudio de mutantes de pérdida o ganancia de función de miembros de la maquinaria de SUMO es esencial para elucidar las funciones en las que está implicado SUMO. Sin embargo, los mutantes de los componentes de la maquinaria de SUMO tienen efectos muy pleiotrópicos que dificultan el estudio de SUMO durante el desarrollo. Mediante conocimientos estructurales de la ruta de conjugación de SUMO, hemos desarrollado una herramienta molecular para inhibir la SUMOilación basada en la disrupción de la interacción molecular entre la enzima activadora y conjugadora de SUMO mediante la expresión del dominio SAE2^{UFDCt}. Durante el ciclo de conjugación de SUMO, la enzima conjugadora de SUMO es reclutada por la enzima activadora mediante interacciones no covalentes mediadas por el dominio de la cisteína catalítica y el dominio UFDCt de SAE2. Mediante ensayos de expresión transitoria en cebolla hemos demostrado la interacción *in vivo* de la enzima conjugadora de SUMO con el dominio SAE2^{UFDCt}, y su reclutamiento hacia el núcleo. El dominio SAE2^{UFDCt} interactúa con la enzima E2 a través del motivo LHEB2, altamente divergente, lo que sugiere que las interacciones E1-E2 son específicos de cada especie. Las plantas que expresan el dominio SAE2^{UFDCt} muestran defectos en el desarrollo característicos de plantas deficientes en la SUMOilación, cómo la floración temprana y defectos en el crecimiento de la planta, confirmando el papel de SAE2^{UFDCt} como inhibidor de la conjugación de SUMO *in vivo*. La localización nuclear de la SAE2^{UFDCt} sugiere que las funciones esenciales de la conjugación de SUMO se producen en el núcleo de la célula. En

general, hemos diseñado y validado una nueva estrategia para inhibir la SUMOilación *in vivo*, que podría ser aplicada para estudiar el papel de SUMO durante etapas de desarrollo específicas en diferentes organismos, ya que las reacciones de conjugación de SUMO están altamente conservadas entre los diferentes reinos.

SUMOylation inhibition *in vivo* is achieved through the expression of the SUMO E1 UFDCt domain.

Laura Castaño-Miquel*, Inês Teixeira*, Josep Seguí, Abraham Mas, Natalia Rodrigo, Gaelle La Verde and L. Maria Lois[1].

Development Program, CSIC- Centro de Investigaciones Científicas.
Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.

*these authors contributed equally to this work

1 Author for correspondence
L.Maria Lois
CSIC- Centro de Investigaciones Científicas
Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain
Tel. +34 93 5636600 ext.3215
Fax. +34 93 5636601
email: maria.lois@cragenomica.es

Running title

In vivo SUMOylation inhibition

Keywords:

SUMO/ inhibitor/ development/ flowering / Alpha Screen

ABSTRACT

Protein modification by SUMO is an essential mechanism that regulates a wide variety of eukaryotic cellular processes. However, pleiotropic defects conferred by strong mutations in SUMO conjugation machinery members hamper the study of SUMO role in development. Taking advantage of the existing structural insights into SUMO conjugation, we have developed a molecular tool for inhibiting SUMO conjugation based on SUMO E1-E2 interactions disruption by the E1 $SAE2^{UFDCt}$ domain. During SUMO conjugation, the E2 conjugating enzyme is recruited to the E1 activating enzyme through two interacting sites located at the E1 $SAE2^{UFDCt}$ and the E1 $SAE2^{Cys}$ domains. We have shown that the $SAE2^{UFDCt}$ domain localizes and recruits the E2 conjugating enzyme to the cell nucleus. The $SAE2^{UFDCt}$ domain interacts with the E2 through a motif, LHEB2, highly divergent across kingdoms, suggesting that E1-E2 interactions are species-specific. Plants expressing the $SAE2^{UFDCt}$ domain displayed developmental defects characteristic of SUMOylation deficient plants in a dose-dependent manner, such as growth defects and early flowering, confirming the role of $SAE2^{UFDCt}$ as a SUMO conjugation inhibitor *in vivo*. The nuclear localization of the $SAE2^{UFDCt}$ suggests that essential functions of SUMO conjugation occur at the cell nucleus. Overall, we have designed and validated a novel strategy to inhibit *in vivo* SUMO conjugation, which could be applied for studying the role of SUMO during specific developmental processes, in agronomic plants and human diseases, since the SUMO conjugation reactions are conserved. In addition, we provide a novel target for developing alternative therapeutic/agronomic drugs.

INTRODUCTION

Many of the plant biological processes regulated by SUMOylation have been uncovered by the analysis of proteases and SUMO E3 ligase mutant plants, which display pleiotropic growth defects and reduced viability (1, 2). Nonetheless, some of these mutations have also been proposed to confer adaptive responses to some stresses, such as salt, drought, resistance to plant viruses and salicylic acid-mediated plant immunity (3-5). As result of the interest raised by the important agronomic traits regulated by SUMO, efforts to study this regulatory system in plants of interest to agriculture have been mainly limited to model plants, such as rice and maize (6, 7). The study of SUMO in other economically relevant plants has been hampered by the lack of molecular tools, such as knock-out mutants. On the other hand, we speculate that the severe growth defects of plants harboring mutations in main components of the SUMOylation machinery, such as *Arabidopsis siz1, mms21 or esd4*, may mask the regulatory role of SUMO in processes occurring late in plant development or, which are overridden by the pleiotropic phenotype. For instance, *siz1* developmental defects are dependent on SA accumulation (8).

Under this scenario, developing tools to achieve specific and/or moderate SUMOylation inhibition would contribute to make significant advances towards an accurate comprehension of SUMO biological role. Considering the universality of the system, the use of these tools could be extended to

other eukaryotes, including mammals. Indeed, the increasing interest in developing SUMOylation specific inhibitors for research and clinical applications is highlighted by previous studies (9-13), which identified natural products or SUMO synthetic derivatives as starting point for chemical optimization and development of therapeutic drugs. The mentioned inhibitors have in common that target single enzymes, such as the E1 activating enzyme (9, 13). But the successful isolation of efficient and specific SUMO conjugation inhibitors will require the identification of novel molecular targets to be added to high-throughput small molecule screens and the development of simple assays to be used in those screens.

SUMO activation is a two-step ATP-dependent reaction catalyzed by the heterodimeric E1 activating enzyme, SAE2/SAE1, which is the first control point in the Ubl selection to enter the conjugation cascade (Fig. S1) (14, 15). SAE2 is structured in four functional domains: adenylation, catalytic cysteine, ubiquitin fold (UFD) and C-terminal domains (16). The E1 activating enzyme small subunit, SAE1, contributes the essential Arg21 to the adenylation domain (17). In *Arabidopsis,* SAE1 has diversified into two isoforms that confer distinct catalytic properties (18), suggesting that SUMO activation could constitute a regulatory step in SUMO conjugation *in vivo*. The adenylation domain is responsible for SUMO recognition and SUMO C-terminus adenylation. After adenylation, the SUMO C-terminal adenylate establishes a thioester bond with the E1 catalytic cysteine. Following thioester bond formation, SUMO can be transferred to the E2 in a reaction that requires E2 recruitment through the $SAE2^{UFD}$ domain (16), although some reports point to a contribution of the $SAE2^{Cys}$ domain (19).

Here, we have shown that $SAE2^{UFDCt}$ functions as a SUMO conjugation inhibitor both *in vitro* and *in vivo* in a dose-dependent manner through a mechanism based on its ability to establish non-covalent interactions with the SUMO E2-conjugating enzyme SCE1. $SAE2^{UFDCt}$ expression is robust and stable through plant generations, indicating that is a reliable strategy to inhibit *in vivo* SUMO conjugation. In addition, our results provide a novel biochemical target and a simplified assay based on Alpha technology to be used in small molecule inhibitors screens, which could result in the identification of novel therapeutic and/or agronomic drugs, in addition to novel research molecular tools.

RESULTS AND DISCUSSION

$SAE2^{UFDCt}$ is essential for SUMO E1 activity and inhibits SUMO conjugation.

In order to develop an alternative and specific SUMO conjugation inhibitor, we have taken advantage of the broad knowledge on SUMO conjugation molecular mechanisms generated in the recent years (20). Specifically, we have exploited the disruption of SUMO E1-activating and E2-conjugating enzyme interactions, which are essential in the SUMO conjugation cascade (16) (Fig. 1A). Interestingly, SAE2 regions involved in SCE1 interactions exhibit a conservation degree ranging from 2-fold to 6-fold lower than the overall conservation displayed among the whole corresponding UFD or Cys domains when *Arabidopsis*, human and yeast orthologs are compared (Fig. S2). This localized divergence suggest that these regions, which we have named LHEB1 and 2 (Low Homology region

involved in E2 Binding 1 and 2), have evolved to optimize cognate interactions in a species-specific manner. Supporting this hypothesis, previous studies performed by us and others showed that the efficiency of the human SUMO conjugation system *in vitro* was dramatically reduced when the human E2-conjugating enzyme was replaced by the *Arabidopsis (21)* or the *Plasmodium (22)* SUMO E2 orthologs. From the E2 side, the region involved in SAE2 binding is better conserved across species and it also participates in SUMO non-covalent interactions (23), which are necessary for polySUMO chain formation (15, 24, 25).

In order to generate a specific and competitive inhibitor of *in vivo* SUMO conjugation, we engineered the $SAE2^{UFDCt}$ domain, which is essential for SUMO conjugation (Fig. 1B) and does not establish multiple interactions during SUMO conjugation, as opposed to the SAE2-interacting domain in the SUMO E2-conjugating enzyme. *In vitro*, the $SAE2^{UFDCt}$ amount present in SUMO conjugation assays directly correlates with the level of SUMO conjugation inhibition (Fig. 1 C and D). Our results indicate that molecular mechanisms involved in SUMO transfer from the E1 to the E2 are conserved in *Arabidopsis* and the $SAE2^{UFDCt}$ domain retains the *in vitro* inhibitory capacity previously shown for the human $SAE2^{UFD}$ domain (16).

$SAE2^{UFDCt}$ colocalizes with SCE1 *in vivo*.

Next, we characterized the capacity of $SAE2^{UFDCt}$ to interact with SCE1. In transient expression experiments in onion cells, SCE1 localized to the nucleus and the cytosol, while the $SAE2^{UFDCt}$ domain localized exclusively to the nucleus, which is consistent with the presence of a nuclear localization signal in SAE2 C-terminal tail (18). When $SAE2^{UFDCt}$ and SCE1 were co-expressed, SCE1 localized exclusively to the nucleus, suggesting that the SCE1 cytosolic fraction was recruited to the nucleus by $SAE2^{UFDCt}$ (Fig 2A). It has been reported that yeast LHEB2 establishes hydrophobic and ionic interactions with Ubc9 (yeast SUMO E2 enzyme), which involve a Leu and two Asp residues, respectively (23). Both interactions play a major role in SUMO transfer from SAE2 to Ubc9 (23), but the low homology between *Arabidopsis* and yeast LHEB2 regions (6% of sequence identity) did not allowed us to identify unequivocally the corresponding functional residues in *Arabidopsis* SAE2. Instead, we analyzed the LHEB2 conservation among plant orthologs (21 from dicots and 5 from monocots) and observed that LHEB2 is highly conserved among plants (Fig. 2B). From the sequence alignment, we identified a consensus sequence, $L_{/I}$-$L_{/V}$-$Y_{/F}$-E-x-G-$D_{/E}$-$D_{/G}$-$L_{/V}$-$E_{/D}$-$E_{/D}$, corresponding to LHEB2 core residues that could be involved in E2 binding according to previous reports (23, 26) (Fig. 3A).

The $SAE2^{UFDCt}$ LHEB2 region has a major role in $SAE2^{UFDCt}$–SCE1 non-covalent interactions.

To analyze the role of some of these conserved residues in E2 binding, we introduced four single mutations into $SAE2^{UFDCt}$, L476A, L477A, D485A and D486A and tested the capacity of the mutant versions to interact with *Arabidopsis* E2 in pull down assays. All $SAE2^{UFDCt}$ mutant forms were impaired in E2 binding, although this defect was accentuated in L476A and D485A mutant forms (Fig. 3B and C). These results are consistent with a prominent role of polar and hydrophobic interactions in

E2 binding. Also, these results show that amino acid residues comprised in the SAE2^{UFDCt} LHEB2 are crucial for establishing SUMO E1-E2 interactions, which are highly divergent across kingdoms. In order to study the affinity of SAE2^{UFDCt} –SCE1 interactions, we have developed a biochemical assay based in Alpha technology (Fig. 3D). From results obtained in competition assays, we have estimated a K_d = 12.8 nM for SAE2^{UFDCt}-SCE1 non-covalent interactions. In addition, these results provide a novel approach to study these bimolecular interactions and constitute the basis for developing a high-throughput small molecule screen, which overcomes the complexity of sumoylation assays previously developed such as electrophoretic mobility shift assays (27), among others. Similar approaches to the described here have successfully identified inhibitors of protein-protein interactions (28).

The SAE2^{UFDCt} domain expression phenocopies the developmental defects displayed by the SUMO E3 ligase mutant *siz1-3* plants.

In order to test the capacity of the SAE2^{UFDCt} domain to inhibit SUMO conjugation *in vivo*, we generated plants expressing *Arabidopsis* SAE2^{UFDCt} domain under the control of the CaMV 35S promoter. We have quantified the endogenous levels of SAE2 and SCE1 in roots and rosettes of 11-day old Col0 plants and determined that there is a molar excess of SCE1 versus SAE2 maintained in both analyzed tissues (8 and 12-fold change in roots and rosettes, respectively) (Fig. S3). Considering this molar excess, an over-accumulation of SAE2^{UFDCt} respect to endogenous SAE2 levels was predicted to be necessary for sequestering endogenous SCE1. Among the obtained transgenic plants, three independent lines expressing different levels of SAE2^{UFDCt} were selected for further characterization (Fig. 4A top). In these plants, accumulation of SUMO conjugates was diminished in direct relationship to SAE2^{UFDCt} expression levels (Fig. 4B). As controls, we analyzed Col0 and *siz1-3* mutant plants, which displayed the highest and the lowest SUMO conjugates accumulation levels, respectively. Remarkably, SCE1 levels were significantly increased in these plants, and this increment was proportional to SAE2^{UFDCt} expression levels. In contrast, SAE2 endogenous levels were not affected (Fig. 4A bottom). It was previously reported that SUMO conjugation deficient *siz1-2* mutant plants displayed an increment of endogenous SCE1 levels and it was proposed as a mechanism to overcome the SUMO conjugation reduction (29). Instead, our results suggest that regulation of endogenous SCE1 levels would involve a more complex molecular mechanism, without excluding a putative compensatory response to SUMO conjugation inhibition. One of the main differences between the SAE2^{UFDCt} expressing plants and *siz1-3* mutant plants is the increased number of contacts that endogenous SCE1 can establish with the LHEB2 domain present in the recombinant SAE2^{UFDCt}. We speculate that SCE1-SAE2^{UFDCt} interactions mediated by the LHEB2 domain could contribute to stabilize SCE1. *In planta*, such mechanism could facilitate the coordination between E1 and E2-levels in order to modulate SUMO conjugation rate. Supporting this hypothesis, we have observed that the molar ratio between endogenous SAE2 and SCE1 is maintained across two analyzed plant tissues, rosette and roots, regardless of the SAE2- and SCE1-level fluctuations among these tissues (Fig. S3).

Moreover, SAE2^{UFDCt} expressing plants displayed developmental alterations previously described in the SUMO E3 ligase *siz1* mutant, such as reduced plant size (Fig. 4C and D) and early flowering (Fig. 4E) (2). The extent of the alterations was consistent with a gradual SUMO conjugation inhibition between the different lines. Previously, strategies aiming to inhibit *in vivo* SUMOylation by expressing a SCE1-inactive mutant version failed as result of the recombinant enzyme instability (21, 30). In contrast, we have shown that the expression of the SAE2^{UFDCt} domain is a reliable and novel approach to inhibit SUMO conjugation *in vivo.*

Most of previous efforts to generate SUMO conjugation inhibitors identified small molecules targeting single enzymes in the conjugation pathway (E1, E2 or E3 enzymes) and their affinity and specificity was rather moderate (9, 12, 13, 27, 31). Alternatively, another study generated semisynthetic protein inhibitors targeting also the E1 activating enzyme that have been tested only *in vitro* (11). Instead, we have disrupted E1-E2 interactions by expressing the SAE2^{UFDCt} domain and shown that this strategy results in an efficient inhibition of SUMO conjugation *in planta*. To our knowledge, this is the first time that is described that the disruption of SUMO E1 –E2 interactions is a valid strategy for inhibiting SUMO conjugation *in vivo*. Disruption of protein-protein interactions potentially offers advantages over single enzyme inhibition related to increased affinity and specificity (32), which in the case of SUMO conjugation is highly relevant in order to identify species-specific inhibitors, and it opens the possibility for therapeutic peptide development.

Overall, we have validated the disruption of SUMO E1 and E2 interactions as a reliable strategy for inhibiting SUMO conjugation *in vivo*. For this purpose we have used the SAE2^{UFDCt} domain as a competitive inhibitor of E1-E2 interactions, showing that the SAE2^{UFDCt} constitutes a novel molecular tool that could be applied not only to study the biological role of SUMOylation in organisms where mutants are not available to perform genetic studies, but also to inhibit SUMO conjugation in a spatiotemporal manner to study specific developmental processes by using inducible or tissue/development-specific promoters. For instance, a recent report has highlighted the interest of identifying SUMO conjugation inhibitors to investigate the essential functions of sumoylation during *Plasmodium falciparum* development (22). Moreover, we have shown that disruption of E1-E2 interactions is a valid target for generating SUMO conjugation inhibitors, which, considering the high divergence presented by the residues involved in these interactions across kingdoms, could be optimized for targeting specific species.

MATERIAL AND METHODS

Plant Material and Growth Conditions. For *in vitro* cultures, seeds were stratified for 3 days, plated on Murashige and Skoog salts (Duchefa), pH 5.7, supplemented with 0.8% BactoAgar (Difco), and transferred to a tissue culture room in a LD photoperiod (16 h light/8 h dark) at 22 °C. For soil cultures, plants were grown in growth chambers in a LD photoperiod at 22ºC.

In Vitro SUMO conjugation. Recombinant proteins were purified as previously described (15). In conjugation assays, we used the C-terminal tail of the Arabidopsis Catalase 3 (419-472) fused to GST, GST:AtCAT3Ct. Reactions were carried out at 37ºC in 25 µL reaction mixtures containing 1 mM ATP, 50 mM NaCl, 20 mM Hepes, pH 7.5, 0.1% Tween 20, 5 mM $MgCl_2$, 0.1 mM DTT, 2 µM SUMO, 0.5 µM AtSAE2/AtSAE1a, 0.5 µM AtSCE1 and 5 µM GST-AtCAT3Ct. After the specified incubation time, reactions were stopped by the addition of protein-loading buffer, incubated at 70 ºC for 10 min, and 10 µL aliquots were resolved by SDS-PAGE. Reaction products were detected by immunoblot analysis with anti-GST polyclonal antibodies (SIGMA, G7781). Luminescence signal generated by ECL Prime assay (GE Healthcare) was captured with a CCD camera (LAS4000, Fujifilm) and quantified with Multigauge software (Fujifilm). Each data point was normalized to the average of all data points obtained from each analyzed membrane in order to remove variability resulting from antibodies incubations and time exposure differences. The normalized values were used to calculate the corresponding slopes (relative luminescence signal versus time). The average slope from at least three independent experiments is shown.

In vitro pull-down assay. 100 µM His:AtSAE2^{UFDCt} or its mutant variants L476A, L477A, D485A and D486A, and 25 µM of AtSCE1 were incubated in 40 µL of binding buffer (50 mM Tris pH 8.0, 150 mM NaCl and 20 mM imidazol) for 1 hour at 4ºC. Next, 10 µL of Ni^{2+}-IMAC-sepharose resin were added to the binding mixture and incubated for 30 minutes at 4ºC. The binding mixture was transferred to micro bio-spin chromatography columns (BIO-RAD, 732-6203) and the resin washed three times with 20 µL of binding buffer and final wash of 40 µL of binding buffer. The proteins bound to the resin were eluted with 20 µL of binding buffer containing 300 mM imidazol. 0.5 µL and 1 µL of the input and eluate fractions, respectively, were separated by SDS-PAGE and subjected to immunoblot analysis with anti-SCE1 antibodies.

Transient Expression of Fluorescent Protein Fusions in Onion Cells. *SAE2^{UFDCt}* and SCE1 were fused in frame to the 3' end of the coding sequences of yellow fluorescent protein (YFP) or cyan fluorescent protein (CFP), respectively, downstream of the 35S constitutive promoter. Onion epidermal cells were bombarded with 5 µg of each DNA construct using a helium biolistic gun (BIO-RAD). Treated epidermal cells were kept in the dark at room temperature for 16 h before analysis by confocal microscopy (Confocal Olympus FV 1000). YFP was excited with a 515-nm argon laser and images collected with a 550- to 630-nm range. CFP was excited with a 405-nm argon laser and images collected in the 460- to 500-nm range. Imaging of YFP and CFP imaging and transmissible light

images collection were performed sequentially. Samples were scanned with the Z-stack mode and image stacks projection was calculated with ImageJ software (33).

Protein interaction assays based on Alpha techonology (Perkin Elmer). His:*SAE2*^{UFDCt} and GST:SCE1 fusion proteins were produced and binding assays were performed according to manufacturer instructions in 25 mM HEPES pH 7.5, 150 mM NaCl, 0.1% BSA and 0.1 mM DTT. Incubations were performed at RT for 1 hour before an additional incubation for 1 hour in the presence of the donor and acceptor beads.

Protein extraction and immunoblot. Anti-SUMO1/2, anti-SAE2 and anti-SCE1 antisera were generated previously(15) Plant tissue was ground in liquid nitrogen and proteins extracted with 50 mM Tris-HCl pH 8, 150 mM NaCl, 0.2% Triton X-100, 1 mM PMSF, 1 µg/mL pepstatin, 1 µg/mL leupeptin, 2 mM *N*-ethylmaleimide, 10 mM iodoacetamide; 5 mM EDTA. 18µg of total protein were resolved under reducing conditions by using SDS polyacrylamide gels and NuPage Novex 4-12% Bis/Tris Gels (Invitrogen). Proteins were transferred onto polyvinylidene difluoride (PVDF) membranes (Millipore), incubated with primary antibody overnight and secondary antibody, peroxidase-conjugated anti-rabbit (GE Healthcare), for 1h at room temperature in TBST (20 mM Tris-HCl at pH 7.6; 20 mM NaCl, 0.1% (v/v) Tween20) supplemented with 3% non-fat dry milk. Peroxidase activity was developed in ECL Plus reagent (GE Healthcare) and chemiluminescence signal captured with LAS-4000 imaging system (Fujifilm).

Accession numbers. Assigned accession numbers for the studied genes are as follows: At5g55160 (SUMO2), At2g21470 (SAE2), At4g24940 (SAE1a), At3g57870 (SCE1).

Acknowledgments

We thank the technical support from members of the Greenhouse and Microscopy facilities at CRAG. We greatly thank Cristina Cañadas for technical support at LML lab. Technical assistance from Rosanna Caliandro is also appreciated. This work was supported by the European Research Council (ERC-2007-StG-205927) and the Spanish Ministry of Science (BIO2008-01495). L.C.M., I.T. and N.R. were supported by research contracts trough the CRAG. A.M. and J.S. were supported by predoctoral fellowships FPU12/05292 and BES-2005-6843, respectively. We also thank the Generalitat de Catalunya (Xarxa de Referència en Biotecnologia and 2009SGR 09626) for substantial support.

Author Contribution

LCM, IT, JS, AM, NR, GLV and LML performed research. LCM, IT and LML designed research. LML wrote the paper. All authors contributed to the analysis of the data and approved the manuscript.

Conflict of Interest

The authors declare that they have no conflict of interest

REFERENCES

1. Ishida T, Yoshimura M, Miura K, & Sugimoto K (2012) MMS21/HPY2 and SIZ1, Two Arabidopsis SUMO E3 Ligases, Have Distinct Functions in Development. *Plos One* 7(10).
2. Lois LM (2010) Diversity of the SUMOylation machinery in plants. *Biochem Soc Trans* 38(Pt 1):60-64.
3. Miura K, Sato A, Ohta M, & Furukawa J (2011) Increased tolerance to salt stress in the phosphate-accumulating *Arabidopsis* mutants *siz1* and *pho2*. *Planta* 234(6):1191-1199.
4. Miura K*, et al.* (2012) SIZ1 deficiency causes reduced stomatal aperture and enhanced drought tolerance via controlling salicylic acid-induced accumulation of reactive oxygen species in Arabidopsis. *The Plant Journal*:n/a-n/a.
5. Yoo CY*, et al.* (2006) SIZ1 SUMO E3 ligase facilitates basal thermotolerance in Arabidopsis independent of salicylic acid. *Plant Physiol.* 142:1548–1558.
6. Wang HD*, et al.* (2011) OsSIZ1 Regulates the Vegetative Growth and Reproductive Development in Rice. *Plant Molecular Biology Reporter* 29(2):411-417.
7. Srilunchang KO, Krohn NG, & Dresselhaus T (2010) DiSUMO-like DSUL is required for nuclei positioning, cell specification and viability during female gametophyte maturation in maize. *Development (Cambridge, England)* 137(2):333-345.
8. Miura K, Lee J, Miura T, & Hasegawa PM (2010) SIZ1 controls cell growth and plant development in Arabidopsis through salicylic acid. *Plant Cell Physiol* 51(1):103-113.
9. Kumar A, Ito A, Hirohama M, Yoshida M, & Zhang KYJ (2013) Identification of Sumoylation Activating Enzyme 1 Inhibitors by Structure-Based Virtual Screening. *Journal of Chemical Information and Modeling* 53(4):809-820.
10. Woo C-H & Abe J-i (2010) SUMO-a post-translational modification with therapeutic potential? *Current Opinion in Pharmacology* 10(2):146-155.
11. Lu X*, et al.* (2010) Designed semisynthetic protein inhibitors of Ub/Ubl E1 activating enzymes. *Journal of the American Chemical Society* 132(6):1748-1749.
12. Fukuda I*, et al.* (2009) Kerriamycin B inhibits protein SUMOylation. *J Antibiot* 62(4):221-224.
13. Fukuda I*, et al.* (2009) Ginkgolic Acid Inhibits Protein SUMOylation by Blocking Formation of the E1-SUMO Intermediate. *Chemistry & Biology* 16(2):133-140.
14. Walden H*, et al.* (2003) The structure of the APPBP1-UBA3-NEDD8-ATP complex reveals the basis for selective ubiquitin-like protein activation by an E1. *Mol Cell* 12(6):1427-1437.
15. Castaño-Miquel L, Seguí J, & Lois LM (2011) Distinctive properties of Arabidopsis SUMO paralogues support the in vivo predominant role of AtSUMO1/2 isoforms. *Biochem J* 436(3):581-590.
16. Lois LM & Lima CD (2005) Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *EMBO J.* 24(3):439-451.
17. Lee I & Schindelin H (2008) Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. *Cell* 134(2):268-278.
18. Castaño-Miquel L*, et al.* (2013) Diversification of SUMO-Activating Enzyme in Arabidopsis: Implications in SUMO Conjugation. *Molecular plant* 6(5):1646-1660.
19. Wang J*, et al.* (2007) The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications. *Mol Cell* 27(2):228-237.
20. Schulman BA (2011) Twists and turns in ubiquitin-like protein conjugation cascades. *Protein Science* 20(12):1941-1954.

21. Lois LM, Lima CD, & Chua NH (2003) Small ubiquitin-like modifier modulates abscisic acid signaling in Arabidopsis. *Plant Cell* 15(6):1347-1359.
22. Reiter K*, et al.* (2013) Identification of Biochemically Distinct Properties of the Small Ubiquitin-related Modifier (SUMO) Conjugation Pathway in Plasmodium falciparum. *Journal of Biological Chemistry* 288(39):27724-27736.
23. Wang J*, et al.* (2010) Crystal Structure of UBA2ufd-Ubc9: Insights into E1-E2 Interactions in SUMO Pathways. *PloS one* 5(12):e15805.
24. Knipscheer P, van Dijk WJ, Olsen JV, Mann M, & Sixma TK (2007) Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. *EMBO J.* 26(11):2797-2807.
25. Capili AD & Lima CD (2007) Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. *J. Mol. Biol.* 369(3):608-618.
26. Wang J*, et al.* (2009) Conformational transition associated with E1-E2 interaction in small ubiquitin-like modifications. *J Biol Chem* 284(30):20340-20348.
27. Kim Yeong S, Nagy K, Keyser S, & Schneekloth Jr John S (2013) An Electrophoretic Mobility Shift Assay Identifies a Mechanistically Unique Inhibitor of Protein Sumoylation. *Chem Biol* 20(4):604-613.
28. Zimmermann G*, et al.* (2013) Small molecule inhibition of the KRAS-PDE[dgr] interaction impairs oncogenic KRAS signalling. *Nature* 497(7451):638-642.
29. Saracco SA, Miller MJ, Kurepa J, & Vierstra RD (2007) Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant Physiol.* 145(1):119-134.
30. Tomanov K*, et al.* (2013) SUMO Conjugating Enzyme with Active Site Mutation Acts as Dominant Negative Inhibitor of SUMO Conjugation in Arabidopsis. *J Integr Plant Biol*.
31. Hirohama M*, et al.* (2013) Spectomycin B1 as a Novel SUMOylation Inhibitor That Directly Binds to SUMO E2. *ACS chemical biology*.
32. Zinzalla G (2013) Understanding and exploiting protein?protein interactions as drug targets. *Understanding and Exploiting Protein?Protein Interactions as Drug Targets,* Future Science Book Series, (Future Science Ltd), pp 2-4.
33. Rasband WS (1997-2009) ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://rsb.info.nih.gov/ij/.

FIGURE LEGENDS

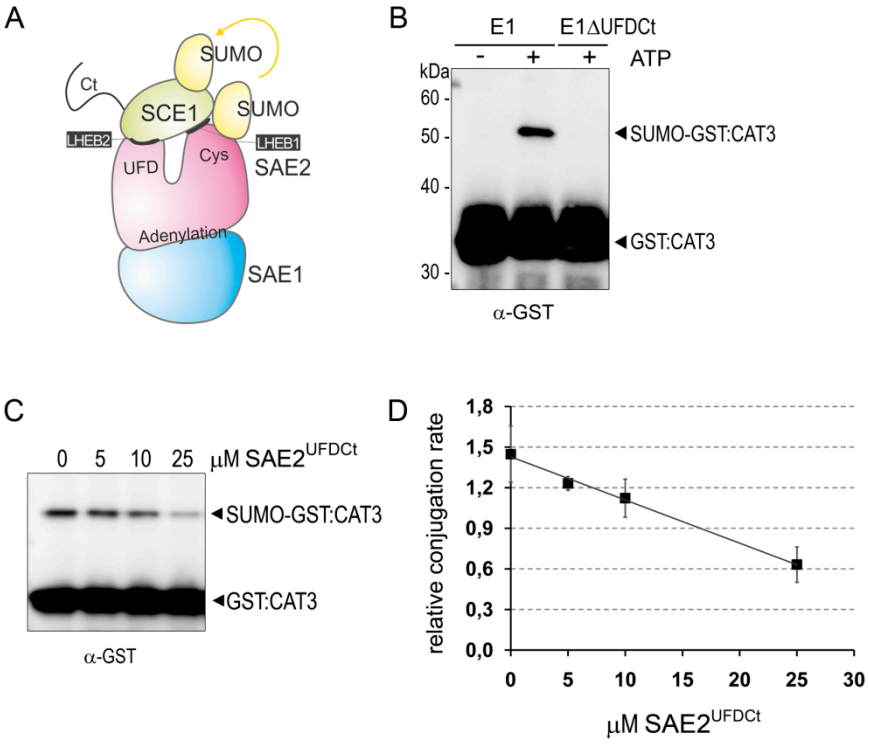


Fig. 1 Engineering SUMO activating enzyme large subunit, SAE2, for inhibiting SUMO conjugation by blocking E1 (SAE2/SAE1) and E2 (SCE1) interactions.

(A) Schematic representation of protein-protein interactions during SUMO transfer from the E1 to the E2.

(B) SAE2^{UFDCt} domain is essential for SUMO conjugation *in vitro*. Sumoylation assays were performed in the presence of Arabidopsis E1 (SAE2/SAE1a) or the deletion mutant E1ΔUFDCt (SAE2 ΔUFDCt/SAE1a), SUMO2, SCE1 and GST:CAT3Ct as substrate. Reactions in the absence of ATP were performed as negative control. Reaction mixtures were incubated at 37 ºC and stopped after 15 minute incubation. Reaction products were resolved by SDS-PAGE and examined by immunoblot analysis with anti-GST antibodies.

(C and D)**.** SAE2^{UFDCt} inhibits SUMO conjugation *in vitro*. Sumoylation assays were performed at 37 ºC in the presence of E1, SUMO2, SCE1, GST:CAT3Ct as a substrate and in the absence or increasing amounts of SAE2^{UFDCt}. Reaction mixtures were stopped after 30 minutes and products were analyzed as in (B). Reactions were performed in quadruplicates and relative GST:CAT3Ct sumoylation quantified. Average values and standard error bars were plotted on the graph (D).

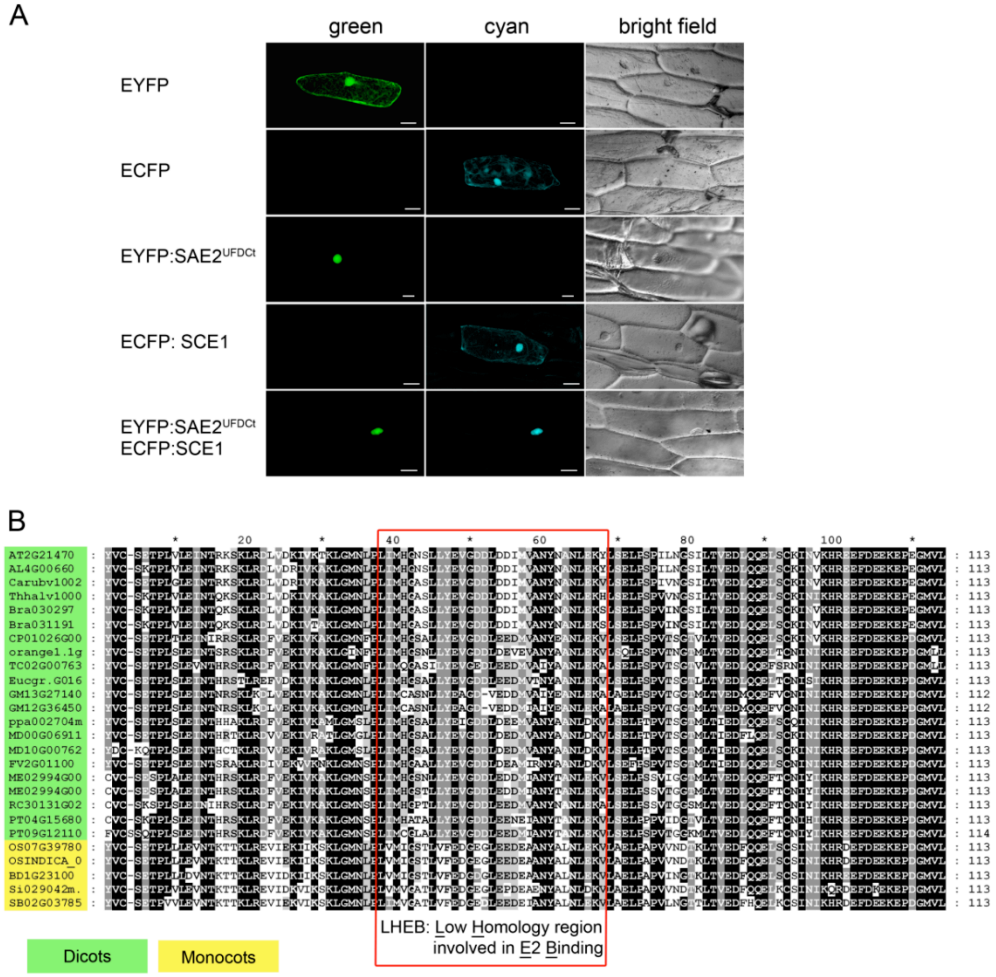


Fig. 2 Analysis of SAE2^{UFDCt} SCE1 interactions.

(A) SAE2^{UFDCt} and SCE1 co-localize in the nucleus of onion cells. SAE2^{UFDCt} fused to EYFP and SCE1 fused to ECFP were transiently expressed in onion epidermal cells, independently or co-expressed. Cells expressing EYFP or ECFP were used as control. Light transmission images of the fluorescent protein expressing cells are shown next to the corresponding fluorescence image. Bars= 50 μm.

(B) Monocots and dicots SAE2^{UFDCt} sequence alignment used for LHEB2 conservation analysis. Black background and white letters corresponds to 100% sequence identity and the LHEB2 region is enclosed in a red box. Gaps in the alignment due to insertions or deletions are indicated by dashed lines. Relative residue numbers are shown to the right of the sequences. Aly, *Arabidopsis lyrata*; Ath, *Arabidopsis thaliana*; Bdi, *Brachypodium distachyon*; Cpa, *Carica papaya*; Csa, *Cucumis sativus*; Gma, *Glycine max*; Mes, *Manihot esculenta*; Mgu, *Mimulus guttatus*; Mtr, *Medicago truncatula*; Osa, *Oryza sativa*; Ppa, *Physcomitrella patens*; Ptr, *Populus trichocarpa*; Rco, *Ricinus communis*; Sbi, *Sorghum bicolor*; Smo, *Selaginella moellendorffii*; Vvi, *Vitis vinifera*; Zma, *Zea mays*.

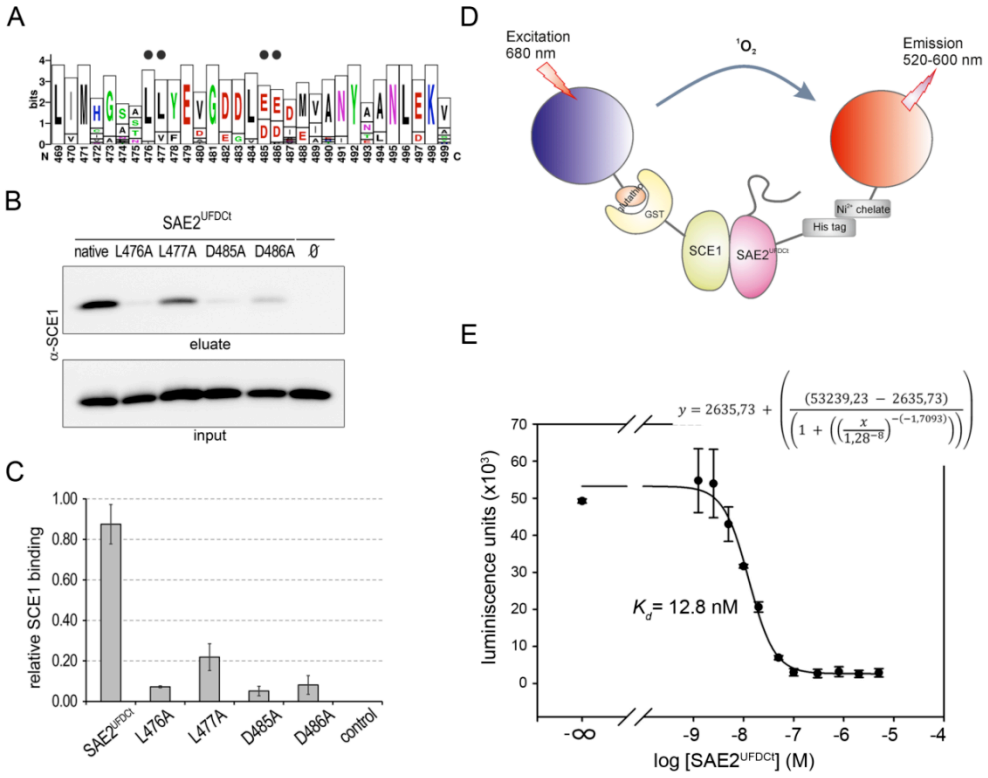


Fig. 3 LHEB2 domain displays a major contribution to SCE1 and SAE2^{UFDCt} interactions.

(A) Graphical representation of plant LHEB2 consensus sequence determined from dicot and monocot SAE2^{UFDCt} sequence alignment. The overall height of the stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each amino acid at that position.

(B) *In vitro* polyHis pull-down assay of Arabidopsis SCE1 using His:SAE2^{UFDCt} or its mutant variants as a bait. Incubations in the absence of the bait were used as negative controls (Ø). Aliquots of input and eluate fractions were resolved by SDS-PAGE and SCE1 levels were analyzed by immunoblotting.

(C) Assays were performed in triplicates and relative SCE1 levels quantified. Average values and standard error bars were plotted on the graph.

(D) Analysis of SCE1-SAE2^{UFDCt} interactions by Alpha technology. GST:SCE1 and His:SAE2^{UFDCt} protein fusions were incubated at room temperature in binding buffer for 1 hour before adding the corresponding donor and acceptor beads. After incubation for one additional hour, luminescent signal was measured with EnVision reader. The schematic representation of the experimental design is shown.

(E) Estimated K_d was calculated from the IC_{50} value obtained in competition assays using the SAE2^{UFDCt} untagged form. Average values of data obtained from two competition assay duplicates were analyzed using a four parameter logistic equation and plotted on the graph. Bars correspond to standard error. The function graph and the estimated K_d are also shown.

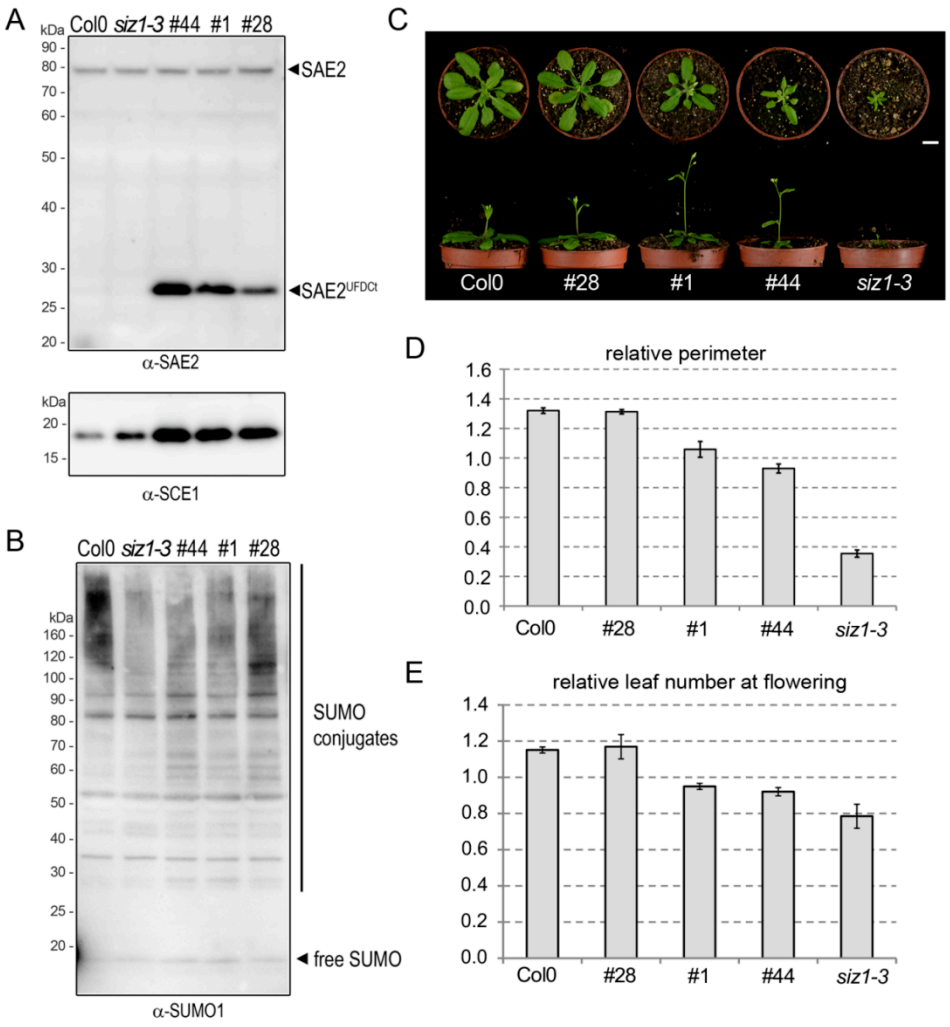


Fig. 4. Effect of SAE2^{UFDCt} expression in endogenous SUMO conjugation and plant development.

(A, B) Effect of SAE2^{UFDCt} expression in SUMO conjugates, SAE2 and SCE1 levels. Total protein extracts from 4-day old seedlings were resolved by SDS-PAGE and examined by immunoblot analysis with (**A**) anti-SAE2, anti-SCE1 and (**B**) anti-SUMO1 antibodies.

(C) Developmental stage of 3-week old plants grown under long day conditions. Scale bar represents 1 cm. Top and lateral views of representative plants are shown.

(D) Rosette perimeter according to ellipse perimeter defined by the three most external leaf tips from each rosette. Average values and standard error bars from relative values obtained in four biological replicates were plotted on the graph.

(E) Rosette leaf number at flowering was scored when the inflorescence had reached 1 cm. Average values and standard errors from relative values obtained in four biological replicates were plotted on the graph.

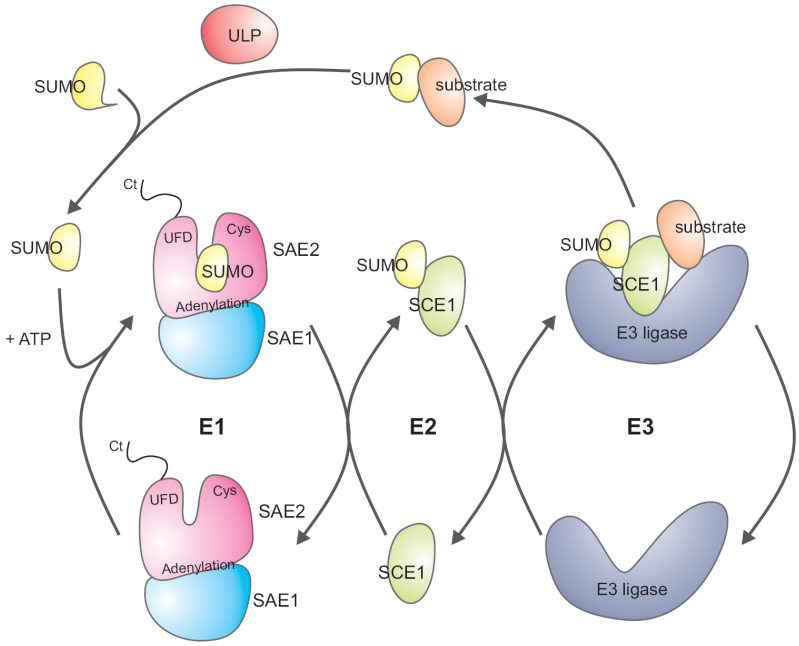


Fig. S1. *SUMO conjugation and deconjugation cycle*. SUMO is synthesized as a precursor that is processed at its C-terminal tail by the specific ULP proteases, releasing a SUMO mature form with a Gly-Gly motif at its C-terminus. SUMO activation is a two-step and ATP-dependent reaction catalyzed by the heterodimeric E1 activating enzyme, SAE1/SAE2. Activation starts with the production of a SUMO-adenylate intermediate that is subsequently transferred to the E1 active cysteine. Next, SUMO is transferred to the E2 conjugating enzyme active cysteine via a transesterification reaction. In a third reaction, SUMO is covalently attached to the target substrate through an isopeptide linkage between SUMO C-terminal glycine and the ε-amino group of the target lysine in the substrate. The target lysine is usually located within the consensus site ΨKxE (Ψ is a large hydrophobic amino acid, K the modified lysine, x any amino acid and E a glutamate acid residue). This final step is facilitated by E3 ligase enzymes that interact both with SUMO charged E2 and the substrate. SUMOylation is a reversible modification and SUMO excision from the substrate is catalyzed by the same class of cysteine proteases involved in the maturation step.

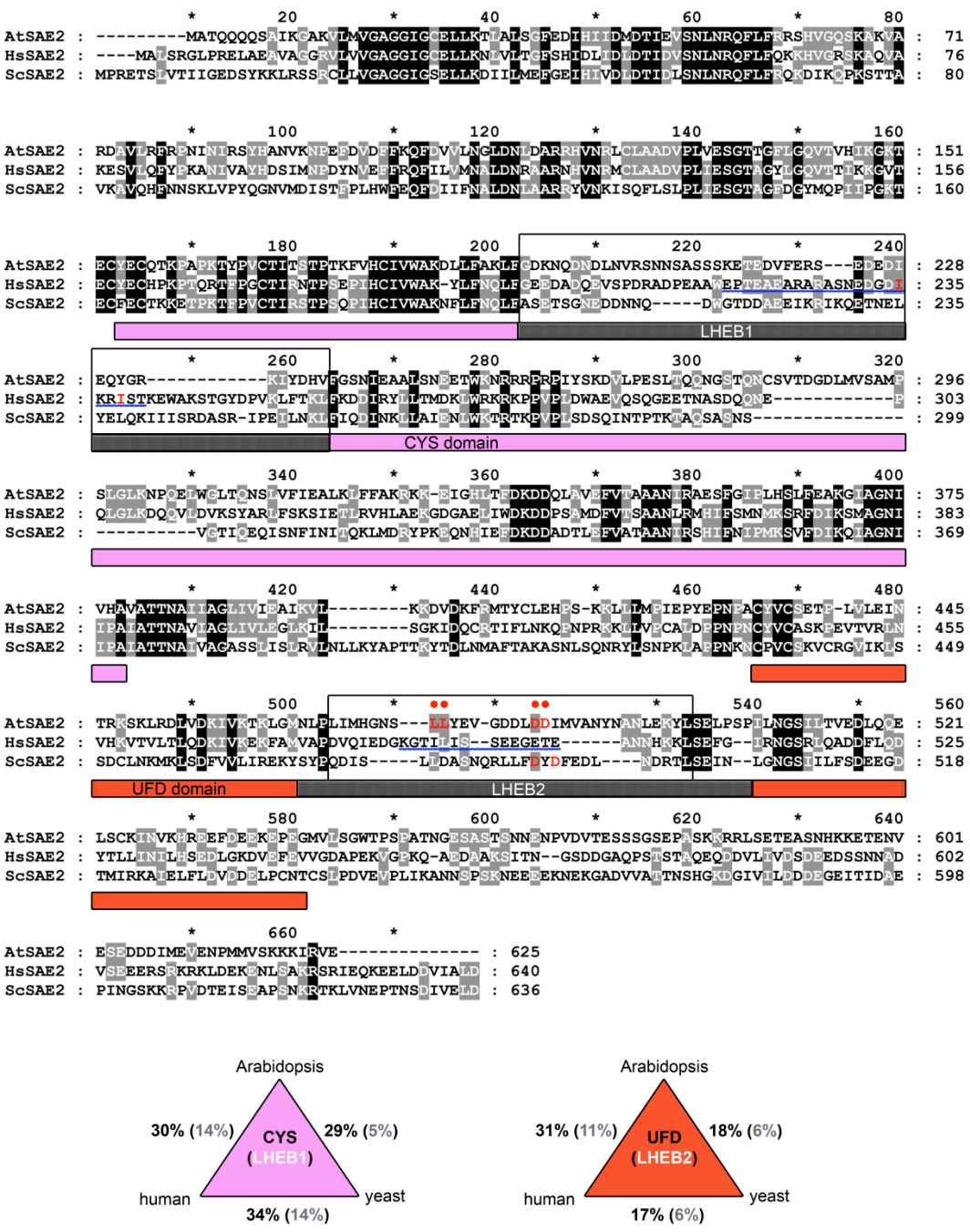


Fig. S2. Arabidopsis, human and yeast SAE2 and SAE1 sequence alignments. Shading colors indicate the following degree of sequence identity: 100% is shown as black background/white letters and more than 60% as grey background/white letters. Gaps in the alignment due to insertions or deletions are indicated by dashed lines. Residue numbers are shown on the right of the sequences. Distribution of SAE2 cysteine and UFD domains are indicated by a pink or red bar below the alignment, respectively. Underlined human SAE2 residues correspond to regions proposed to interact with SUMO E2[1,2]. The contribution in E2-binding of amino acid residues tested by mutagenesis in this study (also indicated by red dots above the sequence) and others are in red[1,3]. Conservation degree of Cys domain and UFD domain among Arabidopsis, human and yeast orthologs is indicated on the

corresponding triangle side. Conservation degree considering only the LHEB1 or the LHBE2 regions is shown in brackets.

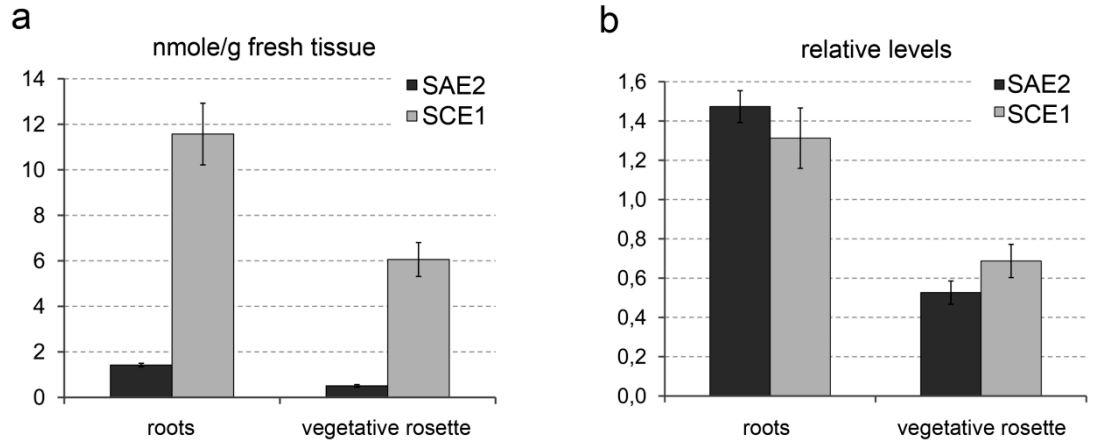


Fig.S3. *SAE2 and SCE1 protein levels in plant tissues.* Plant tissues were collected from 11-day old seedlings grown on half-strength Murashige and Skoog salts supplemented with 1.5% BactoAgar. 15 μg of total protein extracts were resolved by SDS-PAGE and transferred to PVDF membranes. To reduce experimental variability, membranes were divided horizontally, and top and bottom halves were incubated with anti-SAE2 or anti-SCE1 antibodies, respectively. Known amounts of recombinant SAE2 and SCE1 were included in the analysis and used as standards for quantification. Chemiluminescence signal captured with LAS4000 imaging system (Fujifilm) from three biological replicates were quantified and average values and standard error bars were plotted on the graph (A). To facilitate comparison, relative levels are also shown (B). SAE2 and SCE1 values represented are relative to the SAE2 or SCE1 average value, which was calculated considering root and rosette SAE2 or SCE1 absolute levels obtained in (A) from all three biological replicates. Average of relative values and standard error bars from the three biological replicates were plotted on the graph.

1 Wang, J. *et al.* The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications. *Mol Cell* **27**, 228-237 (2007).
2 Wang, J. *et al.* Conformational transition associated with E1-E2 interaction in small ubiquitin-like modifications. *J Biol Chem* **284**, 20340-20348 (2009).
3 Wang, J. *et al.* Crystal Structure of UBA2ufd-Ubc9: Insights into E1-E2 Interactions in SUMO Pathways. *PloS one* **5**, e15805 (2010).
4 Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: A sequence logo generator. *Genome Research* **14**, 1188-1190 (2004).

CAPÍTULO-IV

SUMO conjugation is required for triggering plant defense responses to necrotrophic pathogens

SUMO conjugation is required for triggering plant defense responses to necrotrophic pathogens

Laura Castaño-Miquel[1], Bhagyasree Narayanan[1], Maria Coca[1], Maria Lois[1]

1 Center for Research in Agricultural Genomics CRAG (CSIC-IRTA-UAB-UB), Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.

Resumen. Las plantas son organismos sésiles que están expuestas continuamente al ataque de diversos patógenos y han desarrollado diferentes mecanismos de defensa contra los patógenos. El primer mecanismo de defensa de la planta se basa en el reconocimiento del patógeno que da lugar a la inducción de la resistencia basal. Otra componente importante de esta respuesta es la inducción de la muerte celular programada (HR). Para los patógenos necrótrofos, que obtienen nutrientes para su crecimiento y reproducción de las células muertas del huésped, la HR favorece la infección. Además, se ha descrito que es necesaria una correcta señalización por jasmónico y etileno (JA y ET) para la defensa de la planta frente patógenos necrotróficos. En los últimos años los mecanismos de regulación post-traduccional, como la fosforilación, la ubiquitinización, la nitrosilación, la SUMOilación y la glicosilación se han convertido en factores clave en las respuestas de defensa de las plantas. En las células eucariotas, modificaciones mediadas por SUMO (Small Ubiquitin-related MOdifier) modulan la actividad de la proteína regulando la localización subcelular, estabilidad e interacciones proteína-proteína. Además, se ha establecido una relación entre SUMO y la cascada de señalización activada por ácido salicílico, esencial para la activación de la respuesta sistemática adquirida (SAR) de la planta frente a la infección. Con el fin de estudiar el efecto de SUMO en la respuesta de defensa de la planta frente patógenos necrótrofos, se ha analizado la susceptibilidad de plantas con la SUMOilación afectada frente *Botrytis Cinerea* y *Plectospharaella cucumerina,* patógenos necrótrofos. Los resultados muestran que es esencial un sistema de conjugación de SUMO funcional para la resistencia de las plantas frente patógenos. Sin embargo, se requiere trabajo adicional para entender los mecanismos moleculares que subyacen a esta función. Estos resultados apuntan a la manipulación de la conjugación de SUMO con el fin de establecer un nuevo biomarcador para mejorar la productividad de las plantas.

ABSTRACT

Plants are sessile organisms that are exposed to infection by various pathogens and have developed different defense mechanisms against pathogens. The molecular mechanism underlying activation of plant defense is complex; the first line of defense mechanism is based on the recognition of the pathogen resulting in the induction of basal resistance. Another important plant response component is the induction of the programmed cell death (HR). For necrotrophs pathogens, which obtain nutrients for growth and reproduction from dead cells of the host, HR favours their infection. Moreover, it has been reported that a correct ethylene and jasmonic signalling (JA and ET) pathway is necessary against necrotrophs. Furthermore, in the last years post-translational mechanisms such as phosphorylation, ubiquitination, nitrosylation, glycosylation and sumoylation have become key factors in the resistance of plants against pathogens. In eukaryotic cells, post-translational modifications by SUMO (Small Ubiquitin-related MOdifier) modulate protein activity through several mechanisms, which include regulation of subcellular localization, protein activity and stability, and protein-protein interactions. And it has been described the link between SUMO and SA pathway, which activates systematic acquired resistance (SAR) in plants against biotic stress. In this context, the main objective was the study of the SUMO function in plant defense against necrotrophic pathogens. We analysed the susceptibility of plants with impaired SUMOylation against *Botrytis cinerea* and *Plectospharaella cucumerina*, necrotrophs pathogens. According to the results, we found that a functional SUMO conjugation system is required to display resistance to necrotrophic pathogens. However, additional work is required to understand the mechanisms underlying this role. These results point to SUMO conjugation manipulation as a novel biomarker for improving plant productivity.

I. INTRODUCTION

Plants as sessile organisms are exposed to a broad variety of microbial pathogens and insect herbivores. To protect themselves, plants have developed sophisticated mechanism capable to perceive the pathogen attack and stop the progression of the infection. These surveillance systems are linked to specific programmed defense response. The first lane of defense is triggered via perception of pathogen-associated molecular patterns (PAMPs) by patter-recognition receptors (PRRs) located in the cell surface, called PAMP-triggered immunity (PTI) (Durrant and Dong, 2004; Zipfel, 2008). PTI leads to the induction of basal resistance protecting the plant against a wide range of pathogens. However, pathogens are capable of suppressing basal resistance and infect cell plant. In this case, plants have a second line of defense called the hypersensitive response (HR). The HR is characterized by promoting cell death by reactive oxygen species (ROS) production, in order to limit the access of the pathogen to nutrients and water (Glazebrook, 2005). The HR is typically more pathogen-specific than basal resistance, based on direct or indirect recognition of microbial type III effectors (T3Es) (Mur et al., 2008; Thomma et al., 1998). Once HR has been activated at the local level, plant tissue may become resistant to a wide range of pathogens. This phenomenon is referred as systematic acquired resistance (SAR) and requires SA accumulation that induce expression of pathogen-related (PR) proteins (Durrant and Dong, 2004; Stulemeijer and Joosten, 2008). Other plant defense responses take part in this process, such as the jasmonic and ethylene signalling pathways. The signalling cross talk between these phytohormones is complex and participates on one or another way depending on the pathogen, activating different sets of genes encoding antimicrobial proteins.

Plant pathogens are often divided in two principal groups: biotrophs and necrotrophs. Biotrophs feed on the host alive whereas necrotrophs kill its host and feed on the remains. The plant response is different in front biotrophs and necrotrophs. It has been described that the programed cell death induced by the hypersensitive response promotes resistance in front biotrophs whereas facilitates necrotrophs infection. Furthermore, different defense signaling pathways promotes susceptibility or resistance depending on the plant pathogen. On one hand, the inactivation of the SA pathway leads to increase susceptibility in front biotrophs pathogen, but no effect is observed in front necrotrophs. On the other hand, inactivation of JA response pathway enhanced susceptibility to infection of necrotrophs pathogens. Indicating different hormone-dependent defense pathways, SA pathway against biotrophs whereas the

JA and ET pathway promotes resistance against necrotrophs (Glazebrook, 2005; Thomma et al., 1998).

In the last years, post-translational modification mechanisms have emerged as key players in the plant defense responses to pathogens. The role of phosphorylation, ubiquitination, sumoylation, nitrosylation and glycosylation have been described in plant immunity (Lois, 2010; Stulemeijer and Joosten, 2008). SUMO, *Small Ubiquitin-related MOdifier*, is a small peptide of the ubiquitin family that is conjugated to the lysine residue of the substrate by the sequential function of the E1 activating enzyme (SAE2/SAE1), E2 conjugation enzyme (SCE) and the E3 ligase. This modification is reversible and de-conjugation is carried out by the ubiquitin-like cysteine proteases (ULPs). SUMO has been related to many biological processes, including defense signaling (Lee et al., 2006; van den Burg et al., 2010). The mutant of the SUMO E3 ligase SIZ1 and the sum1-1 amiR-SUMO2 are characterized by high contents of SA resulting in higher expression of PR genes inducing a constitutive systemic-acquired resistance (SAR), which leads to an increased resistance to the bacterial pathogen *Pseudomonas syrinage pv.tomato (Pst)* (Lee et al., 2006; van den Burg et al., 2010). Another evidence for the importance of the SUMOylation system in front pathogen infection comes from the observation that pathogen effectors interact with SUMO machinery components. The *Xanthomonas campestris* effectors XopD and AvrXv4 act as SUMO proteases (Chosed et al., 2007) resulting in the disruption of the SUMO homeostasis in the cell by reducing the amount of SUMO conjugates of the host (Hotson and Mudgett, 2004; Roden et al., 2004). The alteration of the SUMOylation state of the host cell favour infection progression; in viral infections AL1 and REP, essential proteins for viral replication, interacts with SUMO conjugation enzyme, E2, altering the SUMO conjugation state (Castillo et al., 2004; Sanchez-Duran et al., 2011). Moreover, in yeast two-hybrid assays, tomato SUMO (T-SUMO) was identified as an interactor of ethylene-induced xylanase (EIX), which activates ethylene response pathway increasing PR genes accumulation and cell death induction in response to pathogen infection. SUMO-EIX interaction decreases induction of ethylene and cell death (Hanania et al., 1999).

In this context, we aimed to characterize the functional participation of SUMO in plant defense responses to necrotrophic pathogens. For these purpose, the necrotroph fungi *Botrytis Cinerea* and *Plectospharaella cucumerina* were chosen as the model necrotrophic species of study. The *B. cinerea* is an air-borne pathogen considered the second most important infecting more than 200 species (Dean et al., 2012), whereas *P. cucumerina* is a soil-borne emergent pathogen attacking cucurbits (cucumber, melon, pumpkin) (Carlucci et

al., 2012). The resistance to these pathogens is genetically complex and multigenic. To evaluate the role of SUMO conjugation against these necrotrophic pathogens different transgenic and mutant *A.thaliana* lines with altered SUMOyaltion levels were infected. The results obtained suggest that SUMOylation is essential for plant resistance.

II. RESULTS AND DISCUSSION

Plants with impaired SUMOylation exhibit increased sensitivity against fungal pathogens

The role of SUMO in plant defense against necrotrophic fungal pathogens was evaluated with the fungal pathogens *B. cinerea* and *P. cucumerina*. On one hand, *A. thaliana* transgenic plants with increased SUMOylation levels, 35S:*At*SUMO1, were infected. On the other hand, different transgenic lines from 35S:SAE2^{UFDCt} with reduced SUMOylation levels, #1, #28 and #44, were also used. The reduction of SUMO conjugates correlates with the expression level of the SAE2^{UFDCt} domain (#44>#1>#28 SAE2UFDCt expression levels) (see Chapter-III). In addition, mutants of the SIZ1 SUMO E3 ligase (*siz1-3*) and the hereodimeric G-protein (*agb1-1*) were used. The *siz1-3* mutants present the lowest SUMO conjugates accumulation among the used lines and pleiotropic development defects dependent on SA accumulation (Glazebrook, 2005; Miura et al., 2010). The *agb1-1* mutant was used as a positive control of fungal infection due to their susceptibility in front of *P.cucumerina* infection by regulation cell wall composition (Delgado-Cerezo et al. 2012). Four-week old plants were infected with *B.cinerea* and *P.cucumerina,* the progression of the infection was followed by visual inspection. After 7dpi (day-post inoculation), the lines with lowest SUMOylation levels showed and increased suscepitibility to *B.cinerea* whereas no such disease symptoms were observed in the #28, Col-0 and SUMO1 plants (Figure 1A). The disease symptoms caused by *B.cinerea* increased over time, leading to plant decay of #1, #44 and *siz1-3* lines after 15dpi (Figure 1B). In the case of *P.cucumerina* infection, the plants with impaired SUMOylation (#1, #44 and *siz1-3* lines) presented a major susceptibility (Figure 1C). The *siz1-3* and *agb1-1* mutants after 15dpi were mostly death (Figure 1D). A negative correlation was found between SUMOylation levels and the visual disease symptoms caused by the fungal infection (Figure 1E).

Taken together these data suggest that SUMO conjugation is required for resistance against necrotrophic fungus. Fungal infection by *B.cinerea* and *P.cucumerina* resulted in enhanced disease symptoms of plants with diminished SUMOylation, indicating that plant susceptibility correlates with the level of SUMO conjugation inhibition. However, the molecular mechanism involved in SUMO-mediated defense against necrotrophs remains unanswered. Conversely, it is well known that *siz1-3* and *sum1-1 amiR-SUMO2* mutants show enhanced resistance to *PstDC3000* due to SA accumulation, which activates the

immunity innate response by inducing the expression of PR genes (Govrin and Levine, 2000; Lee et al., 2006; van den Burg et al., 2010).

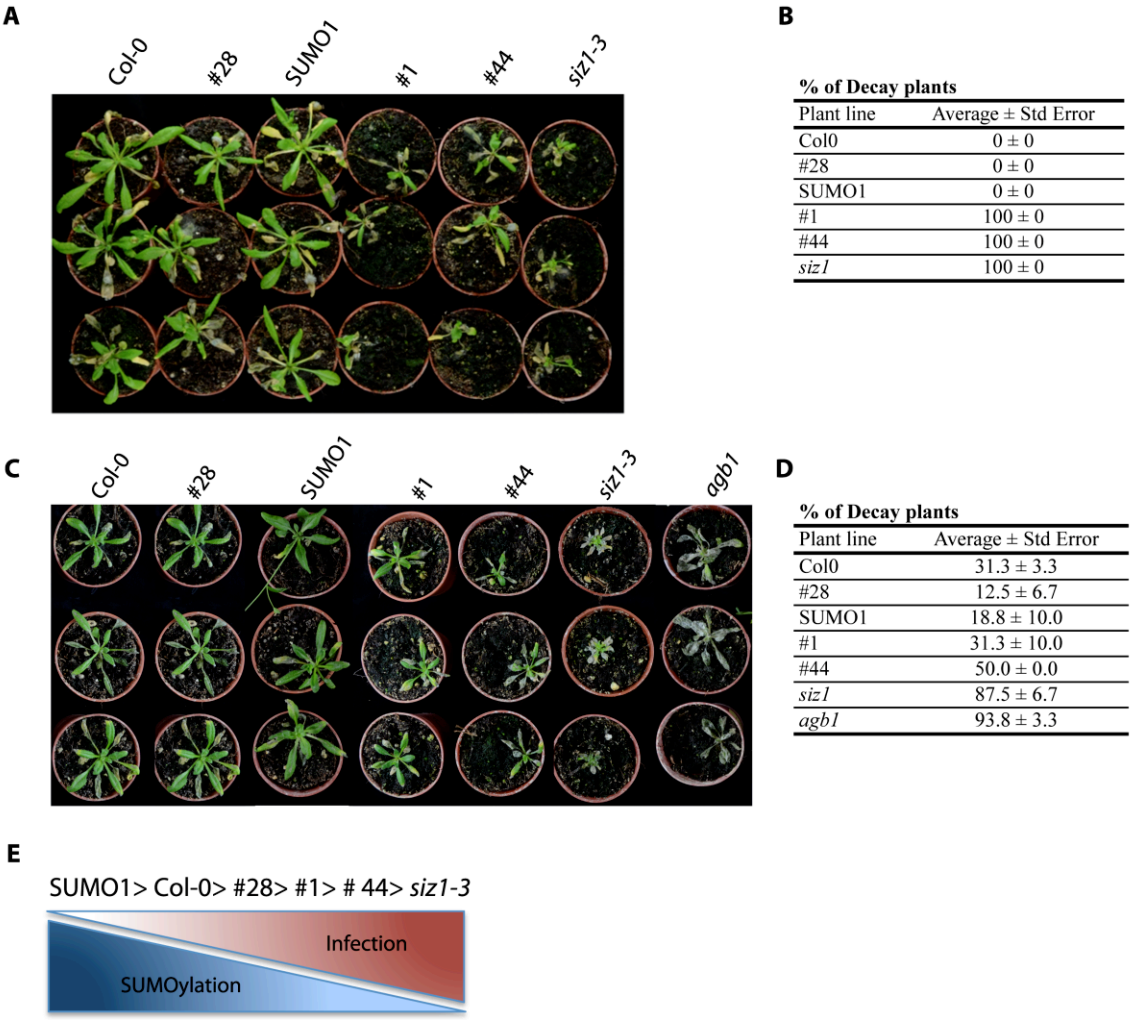


Figure 1. ***SUMOylation is necessary for fungal pathogen resistance.*** ***(A)*** Representative visual disease symptoms of the indicated genotypes at 7dpi with 10^5spores/ml of B.cinerea fungal spore suspensions. ***(B)*** Percentage of decayed plants at 15dpi with 10^5spores/mL of B.cinerea was scored. Average values and standard errors are shown from two independent experiments. ***(C)*** Phenotypical apparence of the indicated genotypes at 15 dpi with 10^5spores/mL of *P.cucumerina* fungal spore suspensions is presented. ***(D)*** Percentage of dead plants a 15 dpi with 10^5spores/mL of *P.cucumerina* is represented. Values correspond to the mean of two independent assays. ***(E)*** Representative scheme of SUMOylation levels and infection progression of the different genotypes. A negative correlation was found between SUMOylation levels and the visual disease symptoms.

PstDC3000 is a bacterial pathogen that triggers SA-dependent defense responses, which play an important role limiting bacterial growth. In contrast, *B.cinerea* and *P.cucumerina* are necrotroph fungal pathogens, and JA and ET- signaling pathways are required for plant resistance (Glazebrook, 2005; Miura et al., 2012). Previous studies reported that jasmonic

acid (JA)-induced PDF1.2 expression and susceptibility to Botrytis cinerea were unaltered in *siz1-2* plants (Lee et al., 2007). In contrast, we have observed that a different *siz1* mutant allele, *siz1-3*, and the SAE2^{UFDCt} expressing plants displayed significant sensitivity to fungal infection. Since fungal infectivity is dependent on growth conditions, we speculate that these difference results may from differences in plant growth environment, fungal spore concentration or a mixture of both.

Taking into account the crosstalk between SUMOylation and JA and ET signaling pathways, future experiments addressing the effect of SUMOylation impairment on JA and ethylene signaling and homeostasis would contribute to understand the molecular basis of this regulation of plant defense responses.

Another important factor in the progression of the pathogen infection are the reactive oxygen species that act at two different levels during the pathogen progression. At early stages, ROS production is involved in the activation of resistance mechanisms in front biotrophs and necrotrophs pathogens (Govrin and Levine, 2000). However, once infection is established increased ROS levels induces HR, which promotes cell death developing resistance in front biotrophs and susceptibility in front necrotrophs. ROS contributes in a different manner depending on the kinetics and context (Glazebrook, 2005). And the same occurs with ROS and SUMO. It is well established that high ROS levels increase SUMO conjugates in response to oxidative stress (Kurepa, 2002). In contrast, low ROS levels results in the decrease on SUMO conjugates by E1-E2 disulfide bond formation in mammalian cells (Bossis and Melchior, 2006), indicating that ROS homeostasis and SUMOylation state in the cell are linked. Alterations in ROS homeostasis could constitute a central player in SUMO modulation in front fungal infection. Additionally, it has been reported that some type III effectors reduce the amount of SUMO conjugates in the host cell (Kim et al., 2008; Roden et al., 2004). In this scenario, an evaluation of SUMOylation dynamics, including SUMO conjugation machinery components, in combination with ROS measurements, in response to fungal infection may be valuable to clarify the molecular mechanism of SUMO action.

In summary, our work identifies a role of SUMO in necrotrophic pathogen defenses, indicating that SUMOylation is required for resistance in front fungal pathogen infection. However, the molecular mechanism remains unclear. Further studies mentioned above should be done to clarify the molecular SUMO implications in plant defenses against pathogen infection. Deeper knowledge may be used for SUMO conjugation manipulation as a novel biomarker for improving plant resistance against necrotrophs infection.

IV. MATERIALS AND METHODS

Preparation of inoculum for fungal infection

Botrytis cinerea, *Plectosphaerella cucumerina* fungal strains were grown on potato dextrose agar plates for 15 days at room temperature. Conidia spores were harvested by adding 5 ml of water, spreading the water using a glass rod to obtain spores, and filtering through Miracloth (Calbiochem). Collected spores were quantified using a counting chamber and a working solution containing the appropriate spore concentration was stored at -80ºC with 30% of glycerol. Before inoculation, 0.02% of tween-20 detergent was added to the prepared spore suspension.

Fungal infection

Arabidopsis thaliana seeds of different lines were grown on soil (Soil ratio =5:1:1; substrate:vermiculita:perlita) in a growth chamber with 12h light/12h dark photoperiod at 22 degree for four weeks. Infections were performed by drop-inoculation, 5 µl spore suspension was put on top of the leaves by using a pipette. The inoculated plants were kept in a plastic box under high humidity for the required period of time.

V. REFERENCES

Bossis, G., and Melchior, F. (2006). Regulation of SUMOylation by Reversible Oxidation of SUMO Conjugating Enzymes. Molecular Cell *21*, 349–357.

Carlucci, A., Raimondo, M.L., Santos, J., and Phillips, A.J.L. (2012). Plectosphaerella species associated with root and collar rots of horticultural crops in southern Italy. Persoonia *28*, 34–48.

Castillo, A.G., Kong, L.J., Hanley-Bowdoin, L., and Bejarano, E.R. (2004). Interaction between a Geminivirus Replication Protein and the Plant Sumoylation System. Journal of Virology.

Chosed, R., Tomchick, D.R., Brautigam, C.A., Mukherjee, S., Negi, V.S., Machius, M., and Orth, K. (2007). Structural analysis of Xanthomonas XopD provides insights into substrate specificity of ubiquitin-like protein proteases. Journal of Biological Chemistry *282*, 6773–6782.

Dean, R., van Kan, J.A.L., Pretorius, Z.A., Hammond-Kosack, K.E., Pietro, A.D., Pietro D Spanu, Rudd, J.J., Dickman, M., Kahmann, R., Ellis, J., et al. (2012). The Top 10 fungal pathogens in molecular plant pathology. Mol. Plant Pathol.

Durrant, W.E., and Dong, X. (2004). Systemic acquired resistance. Annu Rev Phytopathol *42*, 185–209.

Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol *43*, 205–227.

Govrin, E.M., and Levine, A. (2000). The hypersensitive response facilitates plant infection by the necrotrophic pathogen Botrytis cinerea. Curr. Biol. *10*, 751–757.

Hanania, U., Furman-Matarasso, N., Ron, M., and Avni, A. (1999). Isolation of a novel SUMO protein from tomato that suppresses EIX-induced cell death. Plant J. *19*, 533–541.

Hotson, A., and Mudgett, M.B. (2004). Cysteine proteases in phytopathogenic bacteria: identification of plant targets and activation of innate immunity. Current Opinion in Plant Biology *7*, 384–390.

Kim, J.G., Taylor, K.W., Hotson, A., Keegan, M., Schmelz, E.A., and Mudgett, M.B. (2008). XopD SUMO Protease Affects Host Transcription, Promotes Pathogen Growth, and Delays Symptom Development in Xanthomonas-Infected Tomato Leaves. The Plant Cell *20*, 1915–1929.

Kurepa, J. (2002). The Small Ubiquitin-like Modifier (SUMO) Protein Modification System in Arabidopsis. ACCUMULATION OF SUMO1 AND -2 CONJUGATES IS INCREASED BY STRESS. Journal of Biological Chemistry *278*, 6862–6872.

Lee, J., Nam, J., Park, H.C., Na, G., Miura, K., Jin, J.B., Yoo, C.Y., Baek, D., Kim, D.H., Jeong, J.C., et al. (2006). Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase. The Plant Journal *49*, 79–90.

Lois, L.M. (2010). Diversity of the SUMOylation machinery in plants. Biochem. Soc. Trans *38*, 60.

Miura, K., Lee, J., Miura, T., and Hasegawa, P.M. (2010). SIZ1 controls cell growth and plant development in Arabidopsis through salicylic acid. Plant Cell Physiol. *51*, 103–113.

Miura, K., Okamoto, H., Okuma, E., Shiba, H., Kamada, H., Hasegawa, P.M., and Murata, Y. (2012). SIZ1deficiency causes reduced stomatal aperture and enhanced drought tolerance via controlling salicylic acid-induced accumulation of reactive oxygen species in Arabidopsis. The Plant Journal *73*, 91–104.

Mur, L.A.J., Kenton, P., Lloyd, A.J., Ougham, H., and Prats, E. (2008). The hypersensitive response; the centenary is upon us but how much do we know? Journal of Experimental Botany *59*, 501–520.

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

Sanchez-Duran, M.A., Dallas, M.B., Ascencio-Ibanez, J.T., Reyes, M.I., Arroyo-Mateos, M., Ruiz-Albert, J., Hanley-Bowdoin, L., and Bejarano, E.R. (2011). Interaction between Geminivirus Replication Protein and the SUMO-Conjugating Enzyme Is Required for Viral Infection. Journal of Virology *85*, 9789–9800.

Stulemeijer, I.J.E., and Joosten, M.H.A.J. (2008). Post-translational modification of host proteins in pathogen-triggered defence signalling in plants. Mol. Plant Pathol. *9*, 545–560.

Thomma, B.P., Eggermont, K., Penninckx, I.A., Mauch-Mani, B., Vogelsang, R., Cammue, B.P., and Broekaert, W.F. (1998). Separate jasmonate-dependent and salicylate-dependent defense-response pathways in Arabidopsis are essential for resistance to distinct microbial pathogens. Proc. Natl. Acad. Sci. U.S.a. *95*, 15107–15111.

Van den Burg, H.A., Kini, R.K., Schuurink, R.C., and Takken, F.L.W. (2010). Arabidopsis Small Ubiquitin-Like Modifier Paralogs Have Distinct Functions in Development and Defense. The Plant Cell *22*, 1998–2016.

Zipfel, C. (2008). Pattern-recognition receptors in plant innate immunity. Curr. Opin. Immunol. *20*, 10–16.

CAPÍTULO-V

Post-translational modifications of SUMO-Activating Enzyme 2, SAE2, as a regulatory mechanism for modulating SUMO conjugation

Post-translational modifications of SUMO Activating Enzyme 2, SAE2, as a regulatory mechanism for modulating SUMO conjugation

Laura Castaño-Miquel[1], Alfred C.O. Vertegaal[2], L. Maria Lois[1]

1 Center for Research in Agricultural Genomics CRAG (CSIC-IRTA-UAB-UB), Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.
2 Department of Molecular Cell Biology, Leiden Univeristy Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

Resumen. La SUMOilación es una modificación post-traduccional finamente regulada, esencial en eucariotas. La enzima E1 activadora de SUMO, cataliza el primer paso de la cascada enzimática de conjugación de SUMO y es un heterodímero formado por dos subunidades: la subunidad grande, SAE2, y la subunidad pequeña, SAE1. En *Arabidopsis*, mediante estudios bioquímicos se ha descrito que la enzima activadora es capaz de discriminar entre los diferentes parálogos de SUMO y además regula el grado de conjugación de SUMO mediante la diversificación genética de SAE1. Estudios recientes en mamíferos y levadura, han puesto de manifiesto la importancia de diversas modificaciones post-traduccionales que regulan los componentes de la maquinaria de SUMO *in vivo* mediante diferentes mecanismos moleculares que afectan a la localización, la especificidad de sustrato y la actividad de la proteína diana modificada.

En el presente estudio se han identificado dos modificaciones post-traduccionales de SAE2 en *Arabidopsis,* que podrían regular la actividad de la enzima activadora de SUMO *in vivo*. Por un lado, se ha observado el procesamiento proteolítico de ocho aminoácidos del extremo C-terminal de SAE2 durante las etapas finales del desarrollo de semilla. La caracterización funcional de los ocho residuos situados en la cola C-terminal de SAE2 ha puesto de manifiesto su papel como señal de localización nuclear no canónica; indicando que la localización nuclear de SAE2 es dependiente de dos NLS, NLS1 y la señal NLS2 procesable, localizadas en el extremo C-terminal. Mientras que la forma completa de SAE2 presenta localización nuclear, la forma procesada también está presente en el citoplasma. Curiosamente, aunque el dominio C-terminal de SAE2 se encuentra poco conservado entre especies, las secuencias que corresponden a ambas NLSs se encuentran conservadas entre los parálogos de SAE2. Los resultados obtenidos sugieren que esta modificación puede estar

actuando como un mecanismo de regulación de la conjugación de SUMO *in vivo* mediante el control de la disponibilidad del E1 en la célula en etapas específicas de desarrollo y en respuesta a diferentes estreses. Por otro lado, se ha analizado la modificación de SAE2 por SUMO y de otros miembros de la maquinaria del sistema de SUMOilación. Sin embargo, se requieren estudios adicionales con el fin de determinar las consecuencias biológicas y moleculares de estas modificaciones. Modificaciones similares han sido identificadas en el sistema de SUMOilación de levadura y mamíferos, sugiriendo que estas modificaciones podrían constituir un mecanismo conservado de regulación de la actividad, estabilidad y localización de estas enzimas *in vivo*.

ABSTRACT

SUMOylation is an essential post-translational modification in eukaryotic organisms tightly regulated. The SUMO E1-activating enzyme, which controls the first step of the SUMOylation pathway, is an heterodimer which consisting of a large and small subunit called SAE2 and SAE1, respectively. In *Arabidopsis*, it has been described that the E1 regulates SUMO conjugation by selecting SUMO isoforms entering to the SUMOylation pathway and the conjugation rate by SAE1 genetic diversification. Recently studies in mammals and yeast have been focus on the role of diverse post-translational modifications of the SUMO machinery components, which modulate SUMO conjugation *in vivo* by different mechanisms, affecting the localization, target specificity or activity of the modified component.

In this study, we describe two novel post-translational modifications targeting SAE2 in Arabidopsis. On one hand, eight amino acids located at the SAE2 C-terminus are processed during seed maturation by an unknown mechanism. The functional characterization of the eight residues located at the C-terminus tail of SAE2 has evidenced their role as a non-canonical nuclear localization signal, NLS; indicating that SAE2 localization is dependent on two NLS, NLS1 and the processed NLS2, located at the C-terminus tail. Full length SAE2 is located at the nucleus, whereas processed SAE2 is also located at the cytoplasm. Interestingly, regardless of the flexibility and low conservation degree of the SAE2 C-terminal domain, sequences corresponding to both NLSs are conserved among SAE2 paralogs. We postulate that this modification could constitute a regulatory mechanism for SUMO conjugation *in vivo* through regulation of E1 localization under specific development stages or stress conditions. On the other hand, we have identified and analysed the

modification of the members of the SUMOylation system by SUMO, including SAE2, SUMO, SCE1 and SAE1a. Nevertheless further research has to be done to extend our knowledge of the biological and molecular consequences of these post-translational modifications identified. The findings suggest that this modification may be a conserved mechanism to modulate activity, stability and localization of the SUMO machinery components *in vivo,* since similar modifications have been found in mammals and yeast.

I. INTRODUCTION

Protein modification by SUMO (small ubiquitin-related modifier) has emerged as an essential regulatory mechanism in eukaryotes. SUMO belongs to the ubiquitin-like protein family (Ubl), however, in contrast to the main role of ubiquitin tagging proteins for degradation, SUMO modulates protein activity through regulation of subcellular localization, protein stability, protein-protein interactions and by competing with other post-translational modifications (Wilkinson and Henley, 2010). Recent studies have determined that polySUMO chains are substrate of SUMO-targeted Ubiquitin ligases as a novel mechanism for proteolytic control of sumoylated proteins (Sriramachandran and Dohmen, 2014). An enzymatic cascade conserved in all Ubl's mediates SUMO conjugation to specific targets. This enzymatic cascade consists of three main steps: activation, conjugation and ligation. The SUMO E1 activating enzyme is the first control point in the regulatory pathway. The E1 is an heterodimer enzyme formed by a large subunit, SAE2, and a small subunit, SAE1. SAE2 is structured in four functional domains: adenylation, catalytic cysteine, ubiquitin fold (UFD) and the C-terminal domains (Lois and Lima, 2005). In *Arabidopsis*, the E1 regulates SUMO conjugation by selecting SUMO isoforms entering to the conjugation pathway and the conjugation rate by the differential participation of SAE1a/b isoforms in the heterodimer formation (Castano-Miquel et al., 2013; Castaño Miquel et al., 2011).

The Arabidopsis E1 displays nuclear localization like their human and yeast orthologues, consistent with the nuclear enrichment of SUMO targets identified in different studies (Miller et al., 2010; Miller et al., 2013). The SAE2 C-terminal tail containing a classical nuclear localization signal, NLS, is necessary for E1 nuclear localization (Castano-Miquel et al., 2013). Interestingly, in mammals although both E1 subunits have distinct functional NLSs, *in vitro* experiments demonstrated that the E1 large subunit Uba2 NLS is the only one required for the efficient import of the E1 complex into the nucleus (Moutty et al., 2011). Moreover, regulation of *Hs*E1 localization is also dependent on post-translational modification by SUMO at the C-terminal domain, which is required for its nuclear retention (Truong et al., 2012b). During the last years the interplay between SUMO and other post-translational modifications have emerged as an important mechanism to regulate SUMO conjugation at the substrate level and the global SUMO homeostasis at the cellular level. In mammals and yeast, phosphorylation, ubiquitination and SUMOylation of the SUMO machinery components result in different final consequences. In yeast, phosphorylation of SIZ1 modulates its localization facilitating SUMO conjugation to specific substrates (Johnson and

Gupta, 2000). In human and yeast, Ubc9 (SUMO conjugation enzyme, E2) is also SUMOylated. SUMOylation of mammalian Ubc9 promotes changes in target protein specificity (Knipscheer et al., 2008), while in yeast Ubc9 modifications by SUMO regulates negatively septin SUMOylation (Ho et al., 2011) and is required for meiosis (Klug et al., 2013). Overall, several studies have demonstrated the important role of post-translational modifications modulating SUMO conjugation *in vivo* by diverse molecular mechanisms.

In this study, we have identified two post-translational mechanisms targeting SUMO-activating enzyme large subunit, SAE2, in *Arabidopsis*. First, we have determined that the C-terminal eight amino acids of SAE2 are processed by an unknown mechanism. The functional characterization of this modification has elucidated the role of these eight amino acids as a new non-canonical NLS, which is required for exclusive nuclear E1 localization. Whereas non-processed E1 displays exclusive nuclear localization, processed E1 is also present in the cytoplasm. We postulate that this processing could modulate subcellular SUMO conjugation *in vivo* by regulation of E1 localization under specific developmental stages or stress responses. Second, we have analysed the modification of SAE2 by SUMO and we have extend this analysis to other members of the SUMOylation machinery. Previous proteomic studies in plants have identified SUMO1, SAE2, SCE1, SIZ1 and ESD4 as a SUMO target *in vivo*. Interestingly, SAE2 and SCE SUMOylation decrease upon heat stress whereas SIZ1 SUMOylation increases. Suggesting that SUMOylation of the E1-E2-E3 components might be a mechanism to modulate SUMO conjugation under specific conditions (Miller and Vierstra, 2010). By mass spectrometry analysis, we have identified several SUMO attachment sites in SAE2, SAE1a, SCE1 and SUMO. Comparative analysis has shown that some of these sites are conserved in mammals and yeast, suggesting that these modifications may be a conserved mechanism to modulate activity, stability and localization of SUMO machinery components *in vivo*.

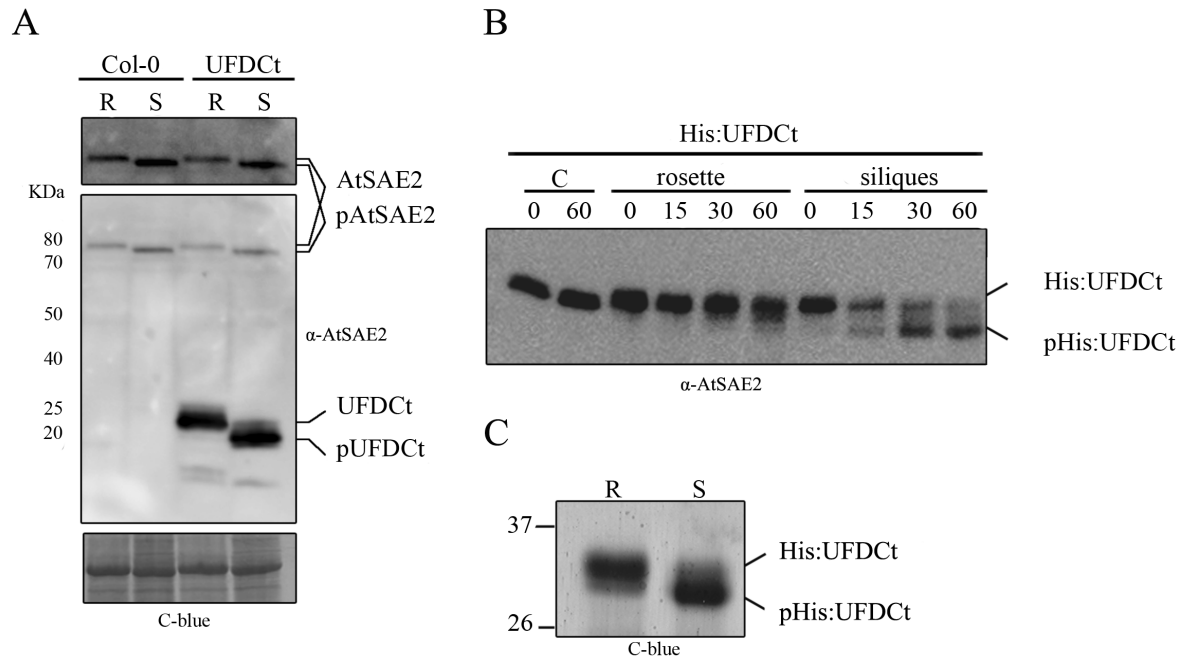
II. RESULTS AND DISCUSSION

SAE2 C-terminus is processed in siliques

We have previously generated plants with impaired SUMOylation system by expressing the SAE2 ubiquitin fold domain, $SAE2^{UFDCt}$, which inhibits the SUMO E1 and E2 interactions required for SUMO conjugation (Castaño-Miquel et al; 2014 unpublished, chapter III). Endogenous SAE2 and recombinant $SAE2^{UFDCt}$ expression levels were evaluated by immunoblotting in rosette and silique tissue extracts from Columbia and $SAE2^{UFDCt}$ expressing plants. Interestingly, different protein mobility of SAE2 and $SAE2^{UFDCt}$ were observed depending on the tissue, indicating the existence of an unknown post-translational modification that would target the UFDCt domain of SAE2 (figure 1A). The UFDCt domain plays an important role in E1 activity; being essential for establishing non-covalent interactions with the E2 through the UFD domain and determining nuclear localization through the C-terminal tail (Castano-Miquel et al., 2013; Lois and Lima, 2005). Taking into account the biological function of these domains this modification could have a regulatory role in the E1 function or localization depending on the domain affected. To discriminate between processing or other post-translational modifications, recombinant His:UFDCt protein was incubated in the presence of silique or rosette protein extracts from Columbia plants. Incubations were performed at 30ºC and aliquots removed at several time points for immunoblot analysis. Figure 1B shows that recombinant UFDCt mobility shifted into a lower molecular weight form in a time-dependent manner in presence of silique protein extracts, while incubation in rosette protein extracts did not resulted mobility shift. Indicating that the different protein mobility observed is due to the proteolytic processing that occurs in silique protein extracts (figure 1B). The incubation reaction was done in presence of different protease inhibitors, which target a wide range of proteases (metaleoproteases, chymiotrypsin, trypsin, lysosomal, aspartic, cysteine and serine proteases). In addition, a bioinformatics analysis of protease recognition sites was performed by PROPSPER software, however the results predicts proteases with broad-spectrum activity against wide range of proteins (Song et al., 2012). Further studies should be done to determine the protease responsible for this modification.

In order to identify the processing site, preparative reactions to generate large amounts of processed and unprocessed $SAE2^{UFDCt}$ forms were performed. Recombinant HIS:UFDCt protein was incubated 60 min in rosette or silique protein extracts from Columbia plants, and

purified taking advantage of their histidine tag. The reaction products were resolved by SDS-PAGE, stained with coomasie blue and analysed by mass spectrometry. The obtained spectra indicated that SAE2^{UFDCt} undergoes processing at the N-terminus of Ser 618, releasing the SAE2 C-terminal eight amino acids (figure 1C).



UFDCt sequence:

```
         *          20         *          40         *
MGSSHHHHHH SSGLVPRGSH MASMTGGQQM GRDPNSSSVD AMSETPLVLE :  50
INTRKSKLRD LVDKIVKTKL GMNLPLIMHG NSLLYEVGDD LDDIMVANYN : 100
ANLEKYLSEL PSPILNGSIL TVEDLQQELS CKINVKHREE FDEEKEPEGM : 150
VLSGWTPSPA TNGESASTSN NENPVDVTES SSGSEPASKK RRLSETEASN : 200
HKKETENVES EDDDIMEVEN PMMVSKKKIR VE                    : 232
                                  NLS1
                    processed C-terminal
```

Figura 1. ***SAE2 C-terminal post-translational modification occurs in a tissue-dependent manner. (A)*** Immunoblot detection of the SAE2 protein. Protein extracts from rosette and silique tissues from Col-0 and SAE2^{UFDCt} expressing plants were resolved by SDS-PAGE and examined by immunoblot analysis with anti-SAE2. ***(B)*** Recombinant His:UFDCt were incubated at 30ºC in the presence of rosette and silique extract from Col-0 plants supplemented with ATP. At the specific time points, aliquots were removed and denatured in the presence of loading buffer. Proteins were resolved in SDS-PAGE and analysed by immunoblotting with anti-SAE2 antibodies. ***(C)*** Unprocessed and processed HIS:UFDCt forms generated by incubation in the presence of rosette and silique extracts were Ni2+-affinity purified and eluted products resolved by SDS-PAGE. Coomassie stained proteins were excised from the gel and analysed by mass spectrometry. SAE2^{UFDCt} sequence showing the identified processing site. Amino acids released by the processing are underlined in red and the canonical NLS is underlined in blue.

SAE2 localization is dependent on two NLSs

The analysis of SAE2 processed sequence showed that is rich in basic amino acids, suggesting that could have functions of nuclear localization signal (Figure 1C). Our group reported that the heterodimeric SUMO E1-activating enzyme localizes preferentially to the nucleus and that this localization was facilitated by a classical NLS located at the C-terminal tail of the SAE2 subunit, which was identified by bioinformatic analyses (Castano-Miquel et al., 2013). Similar studies showed that a nucleo-plasmin type bipartite NLS located in the mammalian E1 large subunit was crucial for E1 nuclear targeting (Moutty et al., 2011). The putative consensus sequence corresponding to bipartite NLS has been described as two basic amino acid clusters separated by 10-12 residues that bind to both importin α binding pockets (Conti et al., 1998; Kosugi et al., 2008). In SAE2 Arabidopsis, 32 amino acids separate the characterized NLS from the eight amino acids processed at the C-terminus, suggesting the presence of two independent NLSs: a classical/canonical NLS and non-canonical NLS, designated as NLS1 and NLS2, respectively. In order to evaluate the putative role of the eight processed amino acids in the control of the E1 localization, a deletion analysis of $SAE2^{UFDCt}$ fused to EYFP C-terminus (enhanced yellow fluorescent protein) was performed and the subcellular localization of these $SAE2^{UFDCt}$ variants determined in onion epidermal cell transient expression experiments. The EYFP: $SAE2^{UFDCt}$ fusion protein localized exclusively in the nucleus whereas the EYFP: $SAE2^{UFDCt}$Δnls2 variants, which has a deletion of the NLS2 sequence, was enriched in the nucleus but is also present at the cytoplasm, as detected by EYFP fluorescence emission (figure 2A). To further evaluate the contribution of NLS1 and NLS2 in the E1 heterodimer nuclear localization, full-length SAE2 or SAE2 variants containing deletions of the NLS1 or NLS2 or NLS1/NLS2 were fused to the EYFP C-terminus. The localization of these SAE2 variants was analysed by co-expression with ECFP:SAE1b in epidermal onion cells. Consistently with the exclusive nuclear localization of the $SAE2^{UFDCt}$, full length SAE2 also displayed an exclusive nuclear localization. In contrast, fluorescence signal was detected in the cytosol when EYFP:SAE2Δnls1 or EYFP:SAE2Δnls2 were expressed, although the ratio nucleus:cytosol distribution was dependent on the preserved NLS. EYFP:SAE2Δnls1 displayed a stronger cytosol localization than EYFP:SAE2Δnls2, suggesting that NLS1 contribution to E1 nuclear localization is quantitatively more relevant than NLS2. Moreover, the cytosolic localization of EYFP:SAE2Δnls1-2 suggests that NLS1 and NLS2 are the main determinants for SAE2

nuclear localization and that both signals are necessary for exclusive localization to the nucleus (figure 2B).

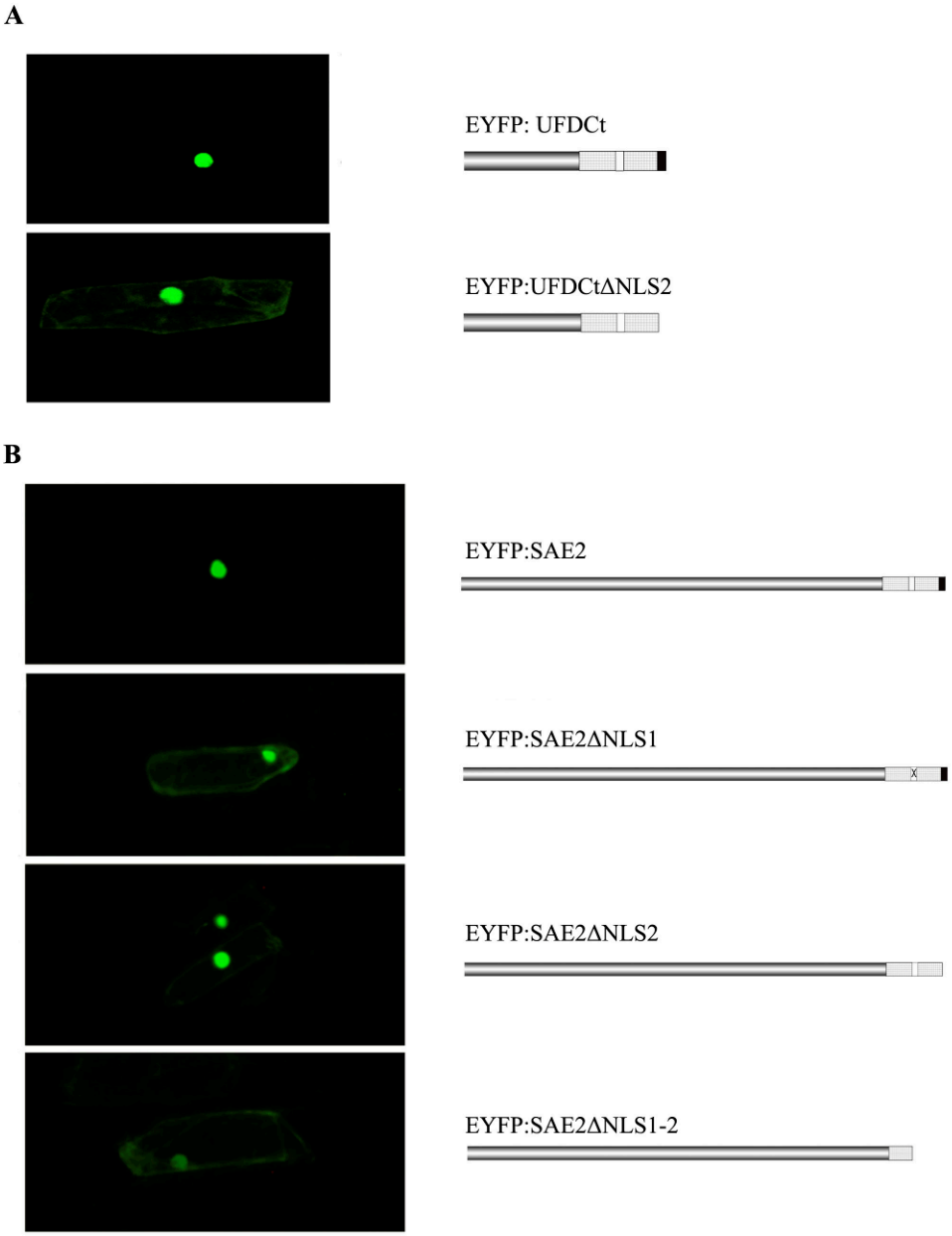


Figure 2. ***Post-translational modification identified as a putative NLS2 (A)*** Transient expression analysis of EYFP:UFDCt and EYFP:UFDCtΔNLS2 in onion epidermal cells is shown ***(B)*** Co-expression of EYFP:SAE2 constructs and ECFP:SAE1b were done. Light transmission images of the EYFP:SAE2 in onion epidermal cells by AixoPhot are shown. Schematic representation of SAE2 and UFDCt constructs generated is shown. The C-terminal domain is represented in light gray whereas the other domains are represented in black gray. The NLS1 and NLS2 are shown in white and black, respectively.

Conservation analysis of SAE2 C-terminal tail

SAE2 homologous sequences were search against Phytozome v.9 database and 47 sequences displaying a conservation ranging from 92% to 29% of sequence identity were retrieved. As previously proposed, the SAE2 UFD and Ct domains are the most divergent domains, specially the C-terminal tail that displays a conservation ranging from 92% to 2% of sequence identity (Figure 3)(Castaño-Miquel et al., 2013). Chlorophyte are the group displaying less sequence conservation to Arabidopsis SAE2 and this effect is accentuated when SAE2 C-terminal tail from these paralogs are compared. SAE2 C-terminal tail from Physcomitrella and Selaginella, which are ancient embriophyte, contain two conserved basic regions that are separated by a short acidic-rich region (figure 3A). In angiosperms, both basic regions are present although they are link by a longer amino acid sequence ranging from 20 to 30 additional amino acids and that also includes the acidic-rich region (figure 3B). These results suggest that differences in SAE2 C-terminal tail length between angiosperms and non-angiosperms are the result of an extension of the region linking both basic-rich regions.

In order to identify conserved functional determinants present in SAE2 C-terminal tail, a new multiple sequence alignment was generated using exclusively SAE2 C-terminal tail sequences from angiosperms and the result is graphically represented by conservation shading, property shading and sequence logo (figure 3B,C). From this analysis, five regions can be identified. Region I corresponds to the consensus sequence GM(V/L)LSGW and it is located at the SAE2 UFD and Ct domains junction. This high conserved motif was used to search for other proteins containing this motif at the Arabidopsis protein database (TAIR 10) and the results indicated that is only present in SAE2. Region II displays a variable extension between sequences ranging from 6 to 19-amino acid long and non specific consensus can be identified. Region III contains a high conserved NG motif. Region IV is also variable, ranging from 21 to 32-amino acid long and only displays a serine-rich region at the N-terminus. Region V constitutes the first basic region, KKR(K/R), corresponding to NLS1 in Arabidopsis SAE2. Region VI is the longest and displays a variable extension ranging from 29 to 47-amino acid long. The C-terminal half of this region displays a conserved acidic/hydrophobic motif. Finally, region VII constitutes the second basic-rich region displaying the conserved motif KK(K/I/Q/R)(R/K) and that corresponds to NLS2 in Arabidopsis SAE2. The conservation degree of these regions among SAE2 paralogs suggests that they could constitute regulatory domains of SAE2 activity *in vivo*, as we have shown to

be the case for NLS1 and NLS2 sequences. Additional experiments would be required to test this hypothesis.

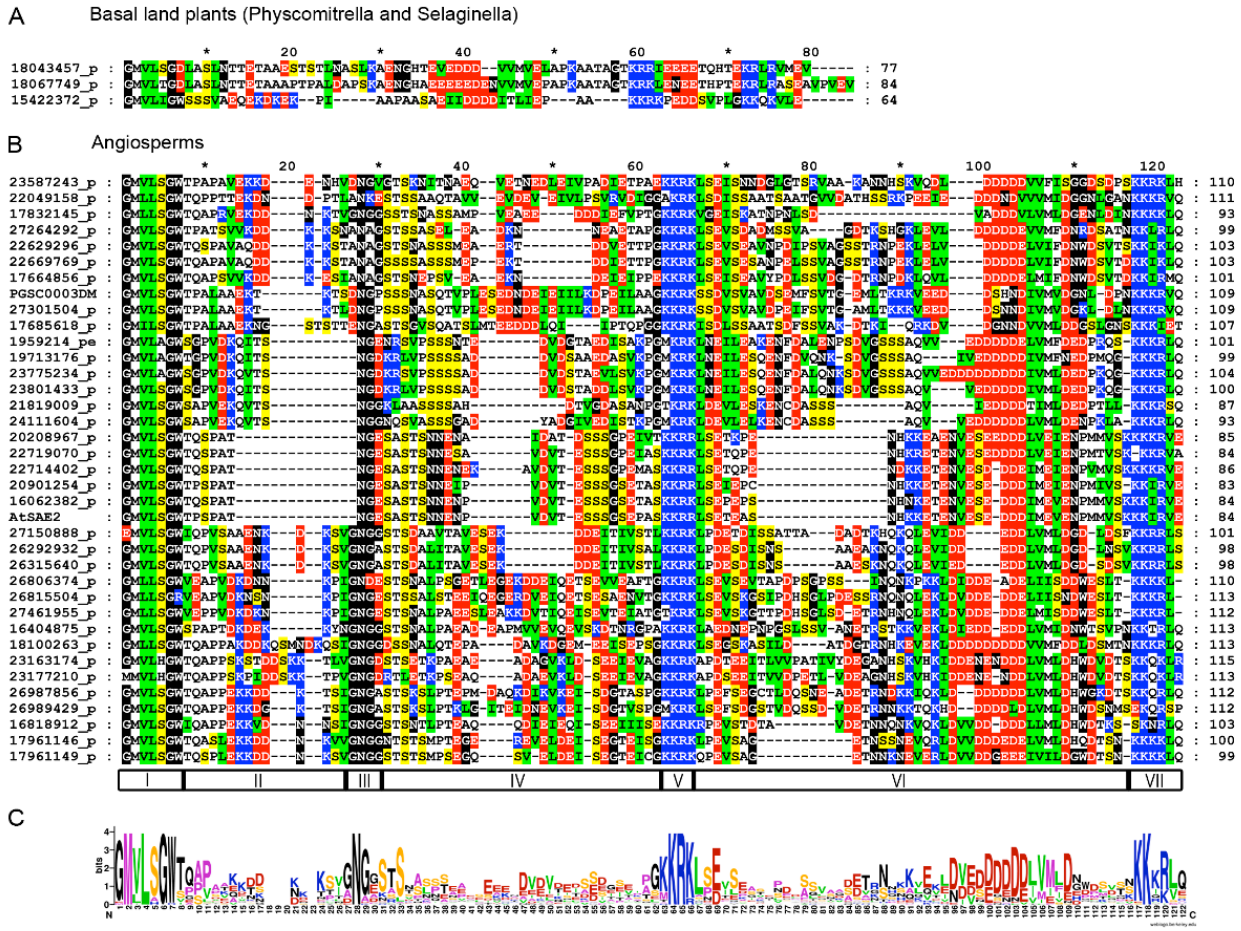


Figure 3. The SAE2 C-terminal domain presents conserved regions, including *NLS1 and NLS2, across Chlorophytes and Angiosperms. (A,B)* Sequence alignments of the SAE2 C-terminal tail from basal land plants and angiosperms are represented; each amino acid is colored according to their chemical properties. *Physcomitrella* and Selaginella presents two basic regions (shown in blue) separated by short acidic region, whereas angiosperms contain a longer linking region (region VI) between the two basic conserved regions, region V and VII identified in *Arabidopsis* as NLS1 and NLS2, respectively. ***(C)*** Sequence logo generated from the alignment of the SAE C-terminal tail sequences from angiosperms. The size and the sorting order of each nucleotide indicate its frequency at each position; the sequence length is showed in numbers above the logo.

These results support our hypothesis of the role of these 8 amino acids processed as a non-canonical nuclear localization signal (NLS2). The subcellular localization of the E1 is regulated by two independent NLSs located at the C-terminus, where the contribution of the NLS1 is stronger but not sufficient for the complete nuclear retention of E1. The NLS2 is essential for the complete nuclear retention of the enzyme. One of the remaining questions to be answer is at which subcellular compartment SAE2 processing takes place. Considering that we have not identified any nuclear exported signals (NES) at SAE2 sequence, and non-

canonical NLS sequences have been described to specifically bind to the minor pocket of importin alpha (Kosugi et al., 2008), our working model is that SAE2 processing may take placed at the cytosol, resulting in a partial translocation to the nucleus

Further studies of the interaction between NLS1/NLS2 and importin alpha are needed to understand the molecular basis of SAE2 nuclear transport. We postulate that this modification has not effect on the activity of SAE2, since the processed NLS2 is located at the C-terminus region, flexible extension that does not interfere in any of the important interactions occurring during the SUMOylation. Finally, the conservation analysis of SAE2 C-terminal tail suggests that, in addition to the conserved NLS1 and NLS2, other identified conserved regions could have additional regulatory roles of SAE2 activity *in vivo*.

We have shown that SAE2 processing is tissue specific, being predominant in seeds. It is well described the important role of SUMO during seed development. SUMO1/2, SAE2 and SCE T-DNA insertion mutants are embryonic lethal; and the mutant of the E3 ligase SIZ1 and HYP1-2 shows a 33,2% and 6,9% aborted seeds, respectively (Liu et al., 2014; Saracco et al., 2007). Although more studies are required, we speculate that SAE2 processing may take place under specific physiological, involving developmental stages and/or stress responses, as a mechanism to modulate SUMO conjugation *in vivo*. In plants and mammals, endogenous SUMO conjugates are enriched into the nucleus of the cell (Bruderer et al., 2011; Miller et al., 2013) and most of the SUMO conjugation machinery members also display preferential nuclear localization (Palancade and Doye, 2007). Nonetheless, several cytoplasmic SUMO substrates have been identified and it remains to be elucidated the molecular mechanism facilitating their modification by SUMO. We propose that SAE2 cytosolic localization mediated by NLS2 processing could favour the SUMOylation of cytosolic substrates under specific conditions. Additional studies should be carried out to elucidate the biological significance of this modification. Currently, transgenic plants overexpressing EYFP:SAE2full, EYFP:SAE2Δnls1, EYFP:SAE2Δnls2 and EYFP:SAE2Δnls1-2 have been done. Their exhaustive study at the physiological and molecular level will reveal the biological relevance of this mechanism of E1 regulation *in vivo*.

SAE2 is SUMO modified.

Recent proteomic studies in plants have identify SAE2 as a SUMO1 substrate (Elrouby and Coupland, 2010; Miller et al., 2010).Interestingly, SAE2 SUMOylation status

decrease during heat stress, suggesting that a non-SUMOylated form of SAE2 is necessary to facilitate the dramatic accumulation of SUMO conjugates in response to heat stress, although the molecular implications of SAE2 SUMOylation has not been studied (Miller and Vierstra, 2010).

To further characterize the biological significance of SAE2 SUMOylation in plants, we aimed to identify the lysine residues in SAE2 that are modified by SUMO based on mass spectrometry analyses. Identification of SUMOylation sites is complex due to the generated pattern of fragmentation of SUMO by MS, in which C-terminal tryptic fragments of SUMOs are frequently very long, hampering the identification of SUMO acceptor lysines in target proteins (Vertegaal, 2011). For this reason, we generated distinct SUMO versions containing artificial trypsin sites next to the C-terminus, that produce more effective footprints for MS to enable identification of SUMO acceptor lysines, as previously published (Knuesel et al., 2005). We generated three SUMO mutant versions, SUMO-H89R, SUMO-Q90R and SUMO-T91R, displaying distinct trypsin footprints as indicated below (figure 4A).

A

SUMO versions	*Trypsin footprints*
SUMO	K + 25 aa
SUMO H89R	K + QTGG
SUMO Q90R	K + TGG
SUMO T91R	K + GG

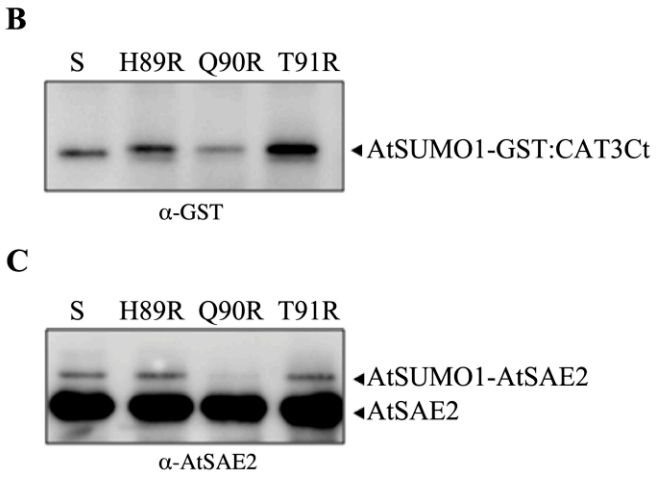


Figure 4. ***HIS-T91R-SUMO1 is competent for conjugation in vitro. (A)*** SUMO variants designed to contain an artificial trypsin digestion site. Trypsin footprints of each SUMO variant are shown. ***(B)*** *In vitro* SUMOylation assays with different SUMO variants were performed in the presence of *At*CAT3Ct as a substrate. Reaction ixtures were incubated at 37ºC during 30 min and analysed by immunoblotting with anti-GST antibodies. **(C)** Conjugation assay as in A was performed; reaction products were analysed by immunobloting with anti-SAE2 antibodies. *At*SUMO-SAE2 formation is showed.

In order to test the conjugation efficiency of the different SUMO1 versions, *in vitro* assays were performed where GST-*At*CAT3Ct was used as SUMO target as previously described (Castaño Miquel et al., 2011). Our results showed that SUMO1-Q90R was conjugated to a lower extend than SUMO1-H89R and SUMO1-T91R, which displayed non-diminished

conjugation efficiencies when compared to native SUMO1 forms (figure 4B). We also tested their capacity to modify SAE2 and we obtained equivalent results to those using *At*CAT3 as SUMO1 substrate (figure 4C). Taking into account the results we proceeded to perform the next experiments using the SUMO1-T91R version, which is not impaired in SUMO conjugation and leads to a di-glicine footprint in the modified substrate.

Generation of SUMOT91R-SAE2 form

To identify SUMO attachment sites in SAE2, we performed *in vitro* conjugation assays using the E1a holoenzyme (SAE2/SAE1a) as substrate. Reaction conditions were scaled-up in order to produce large amounts of SUMO1-T91R modified SAE2 for facilitating MS analysis and two different strategies were carried out to generate the material for mass spectrometry analysis. *In vitro* reactions were performed in the presence of SUMO1-T91R, SCE1 and SAE2/SAE1a and incubated for one hour at 37ºC (figure 5A). On one hand, reaction products were concentrated by trichloroacetic acid (TCA) precipitation and the obtained dry pellet was used for in-solution trypsin digestion and MS analysis. On the other hand, to reduce sample complexity and to enrich the sample in SAE2, unmodified and modified SAE2 were Ni-affinity purified, taking advantage of the His/tag present at the SAE2 N-terminus (figure 5B). The elution fractions were pooled together and resolved by SDS-PAGE electrophoresis (figure 5C). The specific SUMOylated an unSUMOylated bands were excised and subjected to in-gel digestion by trypsin according to standard procedures (Vertegaal et al., 2006) for subsequent mass spectrometry analysis.

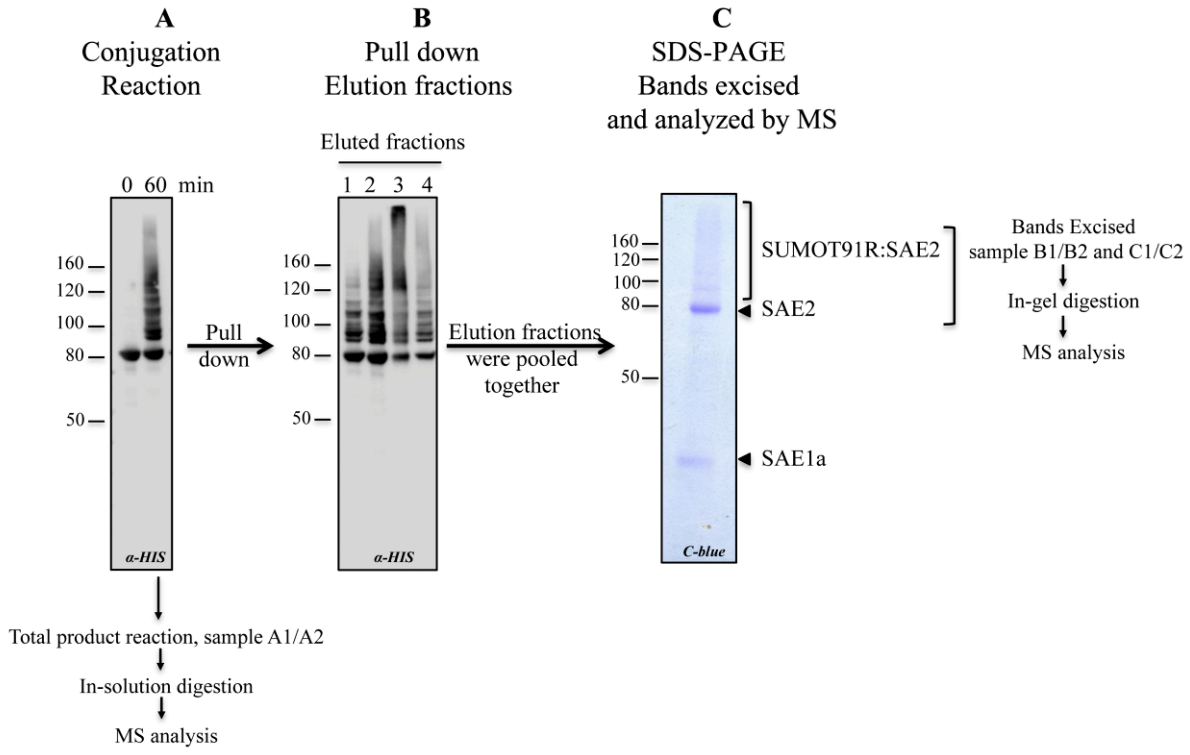


Figure 5. ***Generation of SAE2 SUMOylated forms (A)*** *In vitro* SUMOylation assays were perfomed in presence of *At*SAE2/*At*SAE1a, *At*SUMO1T91R and *At*SCE1. Reaction mixtures were incubated for one hour at 37ºC. Reaction products were analysed by immunoblot with anti-histidine antibodies. The total product obtained in one reaction was precipitated by TCA and subjected to in-solution trypsin digestion followed by MS analysis **(B)** The product reaction of A were purified by IMAC; Elutions of the HIS pull down were analysed by immunobloting. **(C)** Purified elutions were pooled together and loaded in SDS-PAGE gel; bands corresponding to SAE2 unmodified or modified were excised and subjected to in-gel trypsin digestion treatment followed by MS analysis.

Finally, three different products were obtained and duplicate analyses by mass spectrometry were performed yielding six samples: precipitated sample of the total reaction products (A1, A2 samples), SUMO1-T91R-modified SAE2 (B1, B2 samples), and unmodified SAE2 (used as a control, C1, C2 samples) (figure 5 and table supplementary 1).

Identification of the SUMOylation Sites of the SUMO machinery components

The obtained peptides by both approaches were analysed by liquid Chromatography coupled to tandem Mass Spectrometry (LC-MS-MS). Peak assignments of mass spectra allowed the identification of up to 20 SUMO acceptors lysines in SAE2. (table Supplementary 1). Since *in vitro* reactions can lead to unspecific SUMOylation sites, we increased the stringency in the data analysis in order to identify the lysines showing a higher modification degree. Considering the four generated MS data groups, A1, A2 and B1, B2, we selected for further analysis those lysines present in at least three out of the four datasets. We found 2 SUMOylation sites in the C-terminal tail of SAE2 (K595, K619), three in the catalytic cysteine domain (K196, K234, K255) and one in the adenylation domain (K92) (figure 6). The lysine at the position 92 and 196 are found to be located close to the adenylation center of the SAE2 structure and their identification might be due to residual reaction consequence to the proximity with the SUMO-adenylate (Supplemental figure 2A). The lysine at position 234 and 255 are located in the SAE2 cysteine domain and exposed on the surface, suggesting that could be accessible for SUMO conjugating enzyme recognition. The lysine at the position 595 it is located between the two NLSs identified in SAE2. Interestingly, the lysine at position 619 it is found in the new characterized NLS2. Similarly, SUMOylation of human SAE2 also takes place at its C-terminal tail and cysteine domain, which results in nuclear retention and inhibition of SUMO conjugation to target proteins, respectively (Truong et al., 2012b; 2012a). We speculate that SAE2 SUMOylation at the C-terminus in *Arabidopsis* could control nuclear localization of E1 enzyme, as have been described by human SAE2 where SUMOylation at the C-terminus is required for nuclear retention of the enzyme (Truong et al., 2012b). Despite the common localization of the

mapped lysines across the SUMO E1 domains from *Arabidopsis* and *human paralogs*, the exact positions are not conserved (figure supplemental 1A).

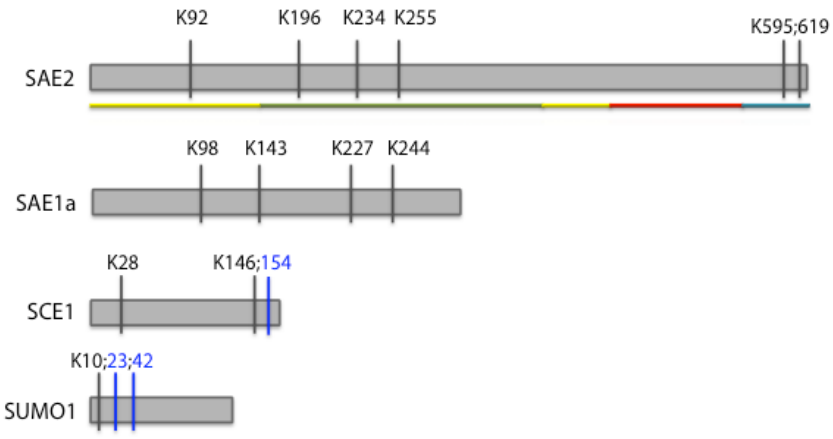


Figure 6. ***Schematic representation of the putative SUMOylation sites identified.*** The most significant modified lysines are shown in the diagram for SAE2, SAE1A, SCE1 and SUMO1. Distribution of SAE2 functional domains, adenylation domain (in yellow), cysteine domain (in green), UFD domain (in red) and C-terminus (in blue) are shown below the schematic protein sequence. Lysine SUMO attachments found *in vivo* by other proteomics studies is shown in blue.

We also identified the SUMO-modified lysines in SUMO1, SCE1 and SAE1a (figure 6 and table supplementary 1). Identified SAE1a sites are close to each other in the tertiary structure and exposed on the surface (Supplemental figure 2A), and conserved in SAE1b isoform. The SAE1a region containing the identified SUMOylation sites is not involved in the interaction with the SAE2 subunit, suggesting that these modifications will not interfere with the E1 heterodimer assembly.

Interestingly, some of the modified lysines found in *At*SUMO1 (K23, K42) and *At*SCE (K154) were previously identified by *in vivo* proteomic analysis (*Miller et al*; 2010) (figure 6 and S2C), supporting our *in vitro* approach to identify biologically relevant SUMOylation sites. The presence of SUMO attachment sites in *At*SUMO1 is consistent with the capacity of this isoform to form polymeric SUMO chains. *At*SUMO2 isoform has also been shown to be able to build polySUMO chains, although only one SUMO attachment site, K10, has been reported to participate in chain formation. Due to the high degree of sequence identity between *At*SUMO1 and *At*SUMO2, it was proposed that SUMO1 could be also SUMOylated at K10 (Colby et al., 2006). Consistent with the conservation mechanism of polySUMO chain formation in both SUMO isoforms, we also identified K10 in SUMO1 as a potential attachment site. These results together with previous showing the SUMOylation of human and yeast SUMO N-terminal tail, *Hs*SUMO1/2 and *Sc*Smt3, and their implication in polymeric SUMO chain formation (Tatham, 2001; Ulrich, 2008) (table 1), support a conserved role of SUMO N-terminus across evolution. Although it is not clear the biological role of polySUMO chain formation, which increase dramatically upon different stress

conditions, it has been described that poly-SUMO chains can act as a ubiquitylation signal targeting the substrates for proteasome-mediated degradation (Tatham et al., 2008).

Identified SUMO acceptor lysines

Organism	Protein	SUMO acceptor lysine	*In vitro*	*In vivo*	Consequences
Human	SUMO2	K11	✓	✓	Polymeric SUMO chain formation
	SUMO3	K11	✓		Polymeric SUMO chain formation
	Uba2	K611, K612, K617, K623, K630	✓		Nuclear retention of the SAE2
		K190, K236, K257, K271, K275	✓		Alterations in the E1 and E2 interaction
	Ubc9	K14	✓	✓	Changes in target recognition
Yeast	Smt3	K11, K15, K19	✓		Polymeric SUMO chain formation
	Ubc9	K153, K157	✓	✓	Decrease SUMOylation of septins
Arabidopsis	SUMO1	K10		✓	NR
		K23,K42	✓	✓	NR
	SUMO2	K10	✓		Polymeric SUMO chain formation
	SAE2	K92, K196, K234, K255, K595, K619	✓		NR
	SAE1a	K98, K143, K227, K244	✓		NR
	SCE1	K154	✓	✓	NR
		K28, K146	✓		NR

Table 1. Identified SUMO acceptor lysines of SUMOs, SUMO E1 and SUMO E2 in Human, Yeast and Arabidopsis. The mapped attachment sites identified by *in vivo* or *in vitro* experiments of SUMO, SUMO E1 and SUMO E2 are shown from human, yeast and *A. thaliana*. In addition, the molecular consequence of this modification is specified in each case. NR: not reported.

In addition, three different SUMOylation sites in SCE1 have been identified located at its N-terminus and C-terminus. (Ho et al., 2011; Knipscheer et al., 2008) (table 1 and figure supplemental 1B and 2B). The identified Lys154 of *At*SCE1 is conserved across *Hs*Ubsc9 and *Sc*Ubc9 and corresponds to the Lys-153 identified as Smt3 attachment site in yeast (figure supplemental 1B). The *Sc*Ubc9 is modified by Smt3 at the C-terminal (Lys-153, Lys 157) and negatively regulates SUMOylation of septins (Ho et al., 2011) (table 1). In contrast, the Lys-153 from *Hs*Ubc9 is not the major SUMOylation site indicating that mammals and yeast use different SUMO attachment sites. In mammals, *Hs*Ubc9 is modified by SUMO at the N-terminus (Lys-14) which regulates target discrimination via gaining or losing affinities, such as impairing SUMO conjugation on RanGAP1 and activating SUMOylation of the transcriptional factor Sp100 (Knipscheer et al., 2008). In addition, the N-terminus of E2s is implicated in the E1 interaction and thioester formation (Bencsath et al., 2002), SUMOylation at the N-terminus of the E2-25K interferes with the ubiquitin E1 interaction (Pichler et al., 2005).

Validation of SUMO acceptors sites

To confirm the identified SUMOylation sites, we used site-directed mutagenesis. For reducing complexity of SAE2 analysis, we used native, K595R or K619R mutant forms of the $SAE2^{UFDCt}$ domain as substrates for testing the identified lysines located at the SAE2 C-terminal tail. To validate the SUMOylation sites at the cysteine domain, site-directed mutagenesis at position 234 and 255 was carried in SAE2ΔCt in order to avoid C-terminal SUMOylation. In the case of SUMO1 and SCE1, site-directed mutagenesis was performed at positions 23, 42 for SUMO and 154 for SCE. *In vitro* SUMO conjugation reactions were performed using the native or the lysine mutant versions of the tested component, except for the reactions using the native or mutant forms of $SAE2^{UFDCt}$ domain as substrates that were performed under standard reactions conditions (figure 7). None of the tested single mutations resulted in abolishment of SUMO1, SCE1 or SCE2 SUMOylation, and only a reduction in SUMOylation was observed in mutants SUMO1K42R, SAE2ΔCtK234R and SAE2UFDCtK595R, suggesting that all tested targets are SUMO modified simultaneously at multiple sites. In the case of SUMO1, it remains to be tested the putative role of K10 as a SUMO attachment site, which was also identified in the proteomic analysis and it would be analogous to SUMO2 SUMOylation at the conserved K10. The putative role of SUMO1K42 as a SUMO acceptor site provides a completely novel molecular mechanism with unknown biological consequences (figure 7A). It is well characterized the capacity of ubiquitin for building chains that are assembled through seven different lysine residues or the amino terminus of ubiquitin, generating eight-homogenous or mixed chains. Instead, it is accepted that SUMO polymeric chains are exclusively connected through a lysine residue located at the N-terminal tail. The identification of SUMO1K42 as an acceptor site for chain formation suggest that polySUMO chains in plants might display higher degree of complexity in contrast to mammalian SUMO paralogs. Four lysines on SCE1 were identified in the MS analysis, although only K154 was unequivocally identified in two replicates and none of them was identified in three replicates. However, the mutagenesis analysis revealed that SCE1K154 is not a major SUMO acceptor (figure 7B). In the case of SAE2, preliminary results showed that K595 and K234 are functional SUMOylation sites although not unique (figure 7C). In general, our results indicate that the different proteins analysed are modified at multiple sites. A more exhaustive analysis of the remaining identified lysine residues by introducing single or multiple mutations are necessary in order to evaluate their relative contribution to SUMOylation of each target. Another possibility is that in absence of the

preferential target lysine, the SUMOylation efficiency of other lysine residues increases under our assay conditions. Increasing the astringency of the SUMOyaltion assays could contribute to evaluate this possibility and uncover the more efficient SUMO attachment lysine residues.

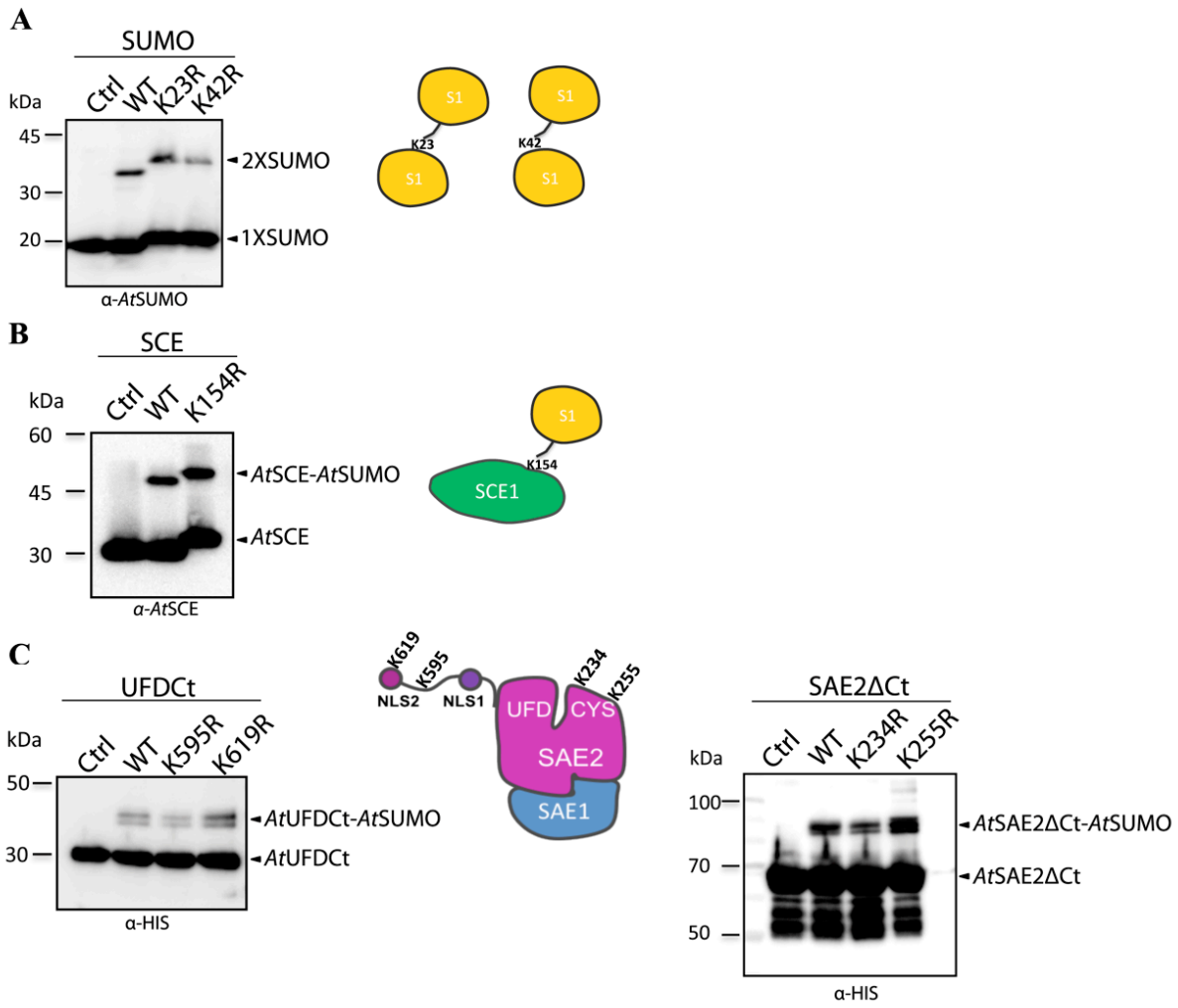


Figure 7. Analysis of putative SUMOylation sites on SUMO1, SCE1 and SAE2. Mutation at the potential modified lysines was performed for SUMO1, SCE1, and SAE2. *In vitro* SUMOylation assays were performed in presence of SUMO1 0,5µM SUMO1, and E1/E2 0,5µM E1 and or the indicated mutant as a substrate 5µM. Reaction mixtures were incubated at 37ºC for 45 min. Reaction products were resolved by SDS-PAGE and examined by immunoblot analysis with the indicated antibodies. Control reactions were performed with all the SUMO machinery components without ATP. (A) SUMOK23R and SUMOK42R was used as a substrate and SUMO conjugation was analysed by immunoblot anti-SUMO.. Schematic representation of the modified lysines of SUMO1 is shown. (B) Contribution of the Lys-154 of SCE1 was analysed, no difference were found between SCE1 and SCE1K154R. Schematic representation of the SCE1-SUMO is shown. (C) Potential SUMOylation sites of SAE2 were examined. The SUMOylation of UFDCtK595R and SAE2ΔCtK234R are reduced. Schematic representation of the distribution of the SAE2 putative SUMO attachment sites studied is shown.

We have identified lysine residues on SAE2, SUMO, SAE1a and SCE1 that undergo SUMO modification. Similar modifications in the SUMO machinery components of yeast and mammals have been previously described, suggesting that these modifications could constitute a conserved regulatory mechanism to control the activity, stability and/or localization of these enzymes *in vivo*. In order to better understand the regulation of SUMO conjugation in plants, a major effort has to be put on the analysis of the functional relevance of these modifications addressing their potential implication in protein activity, stability and/or localization.

Overall, we have shown two distinct post-translational modifications that target SUMO conjugation machinery components, suggesting that SUMOylation is a highly regulated process. On one hand we have identified a novel mechanism for modulating SUMO E1 activating enzyme localization, consisting in the C-terminal processing of the E1-large subunit SAE2, which results in the release of a non-canonical NLS. On the other hand, we have identified SUMO modification of several SUMO conjugation machinery components. Although additional work has to be performed for understanding the biological relevance of these modifications, we speculate that SUMOylation may have similar implications to those described for their mammalian counterparts at the cellular level.

III. MATERIALS AND METHODS

Protein extraction and immunoblot

Plant tissue was ground in liquid nitrogen and proteins extracted with 50 mM Tris-HCl, pH 8, 150 mM NaCl, 0.2% Triton X-100, 1 mM PMSF, 1 μg ml^{-1} pepstatin, 1 μg ml^{-1} leupeptin, 2 mM N-ethylmaleimide, 10 mM iodoacetamide and 5 mM EDTA (extraction buffer). 18 μg of total protein were resolved under reducing conditions by using SDS polyacrylamide gels and NuPage Novex 4–12% Bis/Tris Gels (Invitrogen). Proteins were transferred onto polyvinylidene difluoride (PVDF) membranes (Millipore), incubated with anti-SAE2 antibody overnight followed by incubation with secondary antibody, peroxidase-conjugated anti-rabbit (GE Healthcare), for 1 h at room temperature in TBST (20 mM Tris-HCl, pH 7.6, 20 mM NaCl, 0.1% (v/v) Tween20) supplemented with 3% non-fat dry milk. Peroxidase activity was developed in ECL Plus reagent (GE Healthcare) and chemiluminescence signal captured with the LAS-4000 imaging system (Fujifilm).

Cloning, Expression and Purification of Recombinant Proteins

Proteins for *in vitro* SUMO conjugation assays were cloned into pET28a (Novagene) to encode a N-terminal Histidine tag. The UFDCt and SAE2 constructs for transiently onion expression experiments were cloned in pWEN25 encoding enchanced yellow fluorescent protein (eYFP) for the transient expression experiments. The SUMO versions with artificial trypsin sites at the C-terminal and the constructs to confirm the identified SUMOylation sites were performed by site-directed mutagenesis according to the manufacturer's instructions (QuikChange®, Strategene); the template and the oligonucleotides used for each case are shown below.

Construct	*Primer*	*Oligonucleotides*	*Plasmid template*
SUMO1 T91R	LML284 fw	ATCGATGCGATGCTCCATCAGCGTGGTGGC	PET28a:SUMO1
	LML284 rev	TAGCTACGCTACGAGGTAGTCGCACCACCG	
SUMO1 Q90R	LML 825 fw	GATCGATGCGATGCTCCATCGGACTGGTGGC	
	LML 825 rev	CTAGCTACGCTACGAGGTAGCCTGACCACCG	
SUMO1 H89R	LML 826 fw	ATGAGATCGATGCGATGCTCCGTCAGACTGGTGGC	
	LML826 rev	TACTCTAGCTACGCTACGAGGCAGTCTGACCACCG	
SCE1 K154R	LC117	GGGTGAAGCTGCAGTCCAGGCAGTATCCTGCTCTTGTC	pET28a:SCE1
	LC118	GACAAGAGCAGGATACTGCCTGGACTGCAGCTTCACCC	
SUMO1 K23R	LC121	GCTCACATCAATCTCAAAGTCAGGGGACAGGATGGAAACG	PET28a:SUMO1
	LC122	CGTTTCCATCCTGTCCCCTGACTTTGAGATTGATGTGAGC	
SUMO1 K42R	LC123	GAGAAGCACTCAGCTCAAGAGGCTGATGAATGCTTACTGTG	
	LC124	CACAGTAAGCATTCATCAGCCTCTTGAGCTGAGTGCTTCTC	
SAE2 K255R	LC125	CCAATGAAGAGACATGGAGAAATCGCAGACGACCAAGG	pET28a:SAE2-Ct
	LC126	CCTTGGTCGTCTGCGATTTCTCCATGTCTCTTCATTGG	
SAE2 K234R	LC127	GATATTGAGCAGTATGGTAGGAGGATATATGATCATGTGTTTGG	
	LC128	CCAAACACATGATCATATATCCTCCTACCATACTGCTCAATATC	
SAE2 K619R	LC129	CCATGATGGTGAGCAGAAAGAAGATAAGAGTTGAATAG	pET28a:UFDCt
	LC130	CTATTCAACTCTTATCTTCTTTCTGCTCACCATCATGG	
SAE2 K595R	LC131	CGAGGCCAGTAATCATAGGAAAGAAACAGAGAACGTC	
	LC132	GACGTTCTCTGTTTCTTTCCTATGATTACTGGCCTCG	

The *Arabidopsis* SUMO machinery components generated were purified as described previously (Castaño Miquel et al., 2011) in order to perform the subsequent *in vitro* reactions. In the case of the SUMO variants used for SUMO validation sites, only one step of purification via IMAC-sepharose resin (GE Healthcare) was done. Conversely, for MS analysis all the recombinant proteins were purified via two purification steps; IMAC-sepharose resin followed by size exclusion column. All the histidine tags were removed by thrombin except for SAE2 to allow subsequent purification by IMAC-sepharose resin.

Incubation reaction and purification

350μg of freshly protein extracts from rosette and silique Col-0 plants were incubated with 10μg of recombinant His:UFDCt protein in the presence of the extraction buffer. The control reaction was done by the incubating of the His:UFDCt without any plant extract. Reaction mixtures were incubated during 60 min at 30ºC. After the specified incubated times, reactions were stopped by the addition of protein loading buffer and boiled for 10 min. 15-fold dilution of the result sample was loaded and resolved in 12% SDS-PAGE followed by immunoblot analysis anti-polyHIS monoclonal antibodies (Sigma, H1020). Incubation scale-up experiments were done in order to purify HIS:UFDCt and perform subsequent MS analysis. Scale-up conditions: 100μg of HIS:UFDCt incubated with 3.5mg of rosette or siliques protein extracts from Col-0 plants during 60 min at 30ºC. After the specified incubation time HIS:UFDCt was purified using their histidine tag via IMAC-sepharose (GE Healthcare) according to manucfacter's conditions. Representative elutions of the purified recombinant protein in each condition, incubated with rosette or silique protein extracts, were polled together and concentrated until 40μL using microcon® centrifugal filters. The concentrated products were loaded and resolved by 12% SDS-PAGE gel and stained with Coomassie blue. Bands were excised and subjected to mass spectrometry analysis.

Transient Expression of Fluorescent Protein Fusions in Onion Cells

To analyse the EYFP localization of p35S:EYFP-SAE2 and p35S:EYFP-UFDCt constructs, onion epidermal cells were bombarded with 5μg of each DNA construct using a helium biolistic gun (BIO-RAD). Treated epidermal cells were kept in the dark at room temperature for 16 h before analysis by confocal or AixoPhot microscopy. For YFP, the excitation wavelength was 515-nm argon laser and fluorescence was collected in the range of 550–630-nm. CFP was excited with a 405-nm argon laser and images collected in the 460–500-nm range. An FV 1000 confocal microscopy (Olympus) was used for the analysis of UFDCt and UFDCtΔNLS2 localization, samples were scanned with the Z-stack mode and image stacks

projection was calculated with ImageJ software. In the case of SAE2, SAEΔNLS1, SAE2ΔNLS2 and SAE2ΔNLS1-2 the analysis of protein localization was performed by AixoPhot microscopy.

In vitro SUMO conjugation Assays

In order to validate the activity of the SUMO variants generated, *in vitro* SUMO conjugation assays were done using GST:*At*Cat3Ct and HIS:*At*SAE2 as a substrate. Reaction was carried out at 37ºC for 30 min in 25µL reaction mixture volumes containing 1 mM ATP, 50 mM NaCl, 20 mM Hepes (pH 7.5), 0.1 % Tween 20, 5 mM $MgCl_2$ 0.1 mM DTT (dithiothreitol), 2 µM SUMO variants, 0.5 µM AtSAE2/AtSAE1a, 0.5 µM *At*SCE1 and 5 µM of the substrate. Same conditions reaction was done for the validation of SUMOyaltion sites of SUMO1, SCE1 and SAE2 using the specific mutated forms as a substrate. The reactions were done using the native or the mutated version of the tested component, excluding the native or mutant $SAE2^{UFDCt}$ domain as a substrate that was performed with all the SUMO machinery components. The generation of SAE2 SUMOylated forms was carried out at 37ºC for 1 hour in 400µL of reaction mixture containing 2mM ATP, 50mM NaCl, 20mM Hepes (pH 7.5), 0.1 % Tween 20, 5mM MgCl2 , 0.1mM DTT (dithiothreitol), 10µM SUMOT91R, 2.5µM AtSAE2/AtSAE1a and 2.5µM AtSCE1.

The *in vitro* SUMO conjugation reactions were stopped at the specific time by the addition of protein-loading buffer, boiled for 5 min and 10µL aliquots were resolved by SDS-PAGE and analysed by immunoblotting with anti-GST polyclonal antibody (Sigma, G7781), poly-HIS monoclonal antibody (Sigma, H1020), anti-SUMO and anti-SCE polyclonal antibody as indicated in each panel reaction.

Pull down isolation of SAE2 SUMOylated form

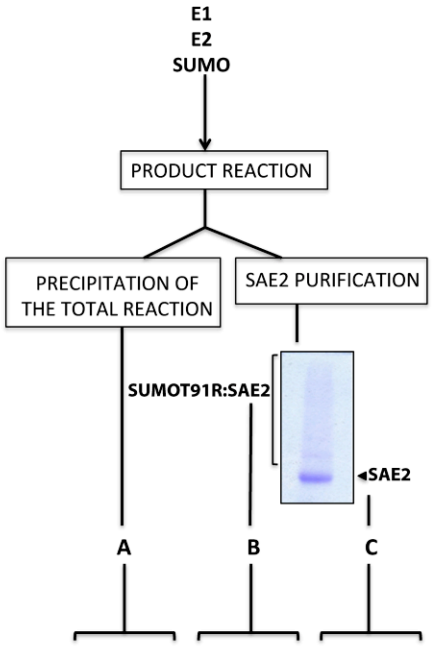
Reaction for SAE2-SUMOT91R formation was performed by duplicate as indicated above. Product reactions were incubated in batch with 50µL of Ni^{2+} -IMAC Sepharose resin at 4ºC for 30 min in the binding buffer (20mM TRIS-HCl pH8.0, 150mM NaCl and 20mM Imidazole pH8.0). The binding mixture was washed four times with 50µL of binding buffer. The proteins bounds to the resin were eluted with 50µL of elution buffer (20mM TRIS-HCl pH8.0, 150mM NaCl and 250mM Imidazole pH8.0). Eluted fractions were separated by SDS-PAGE and subjected to immunoblot analysis with anti-*At*SAE2 antibodies.

Mass Spectrometry Analysis

Purified proteins were size-separated by SDS-PAGE, in-gel digested, ex-tracted, desalted, concentrated, and analysed by mass spectrometry using an EASY-nLC system (Proxeon)

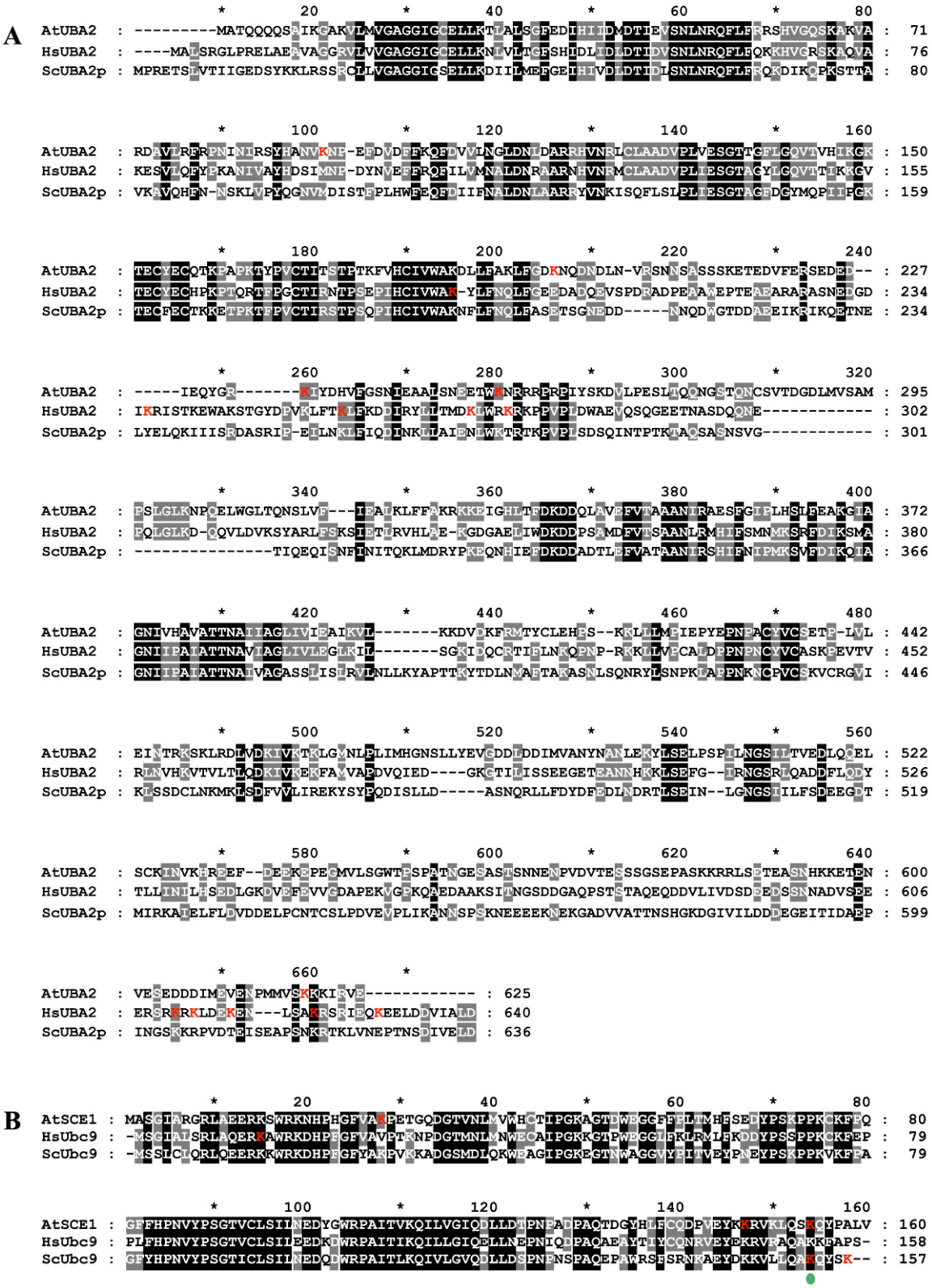
connected to the Q Exactive (both from Thermo Fisher Scientific) through a nanoelectrospray ion source. Raw mass spectrometry (MS) files were processed with the MaxQuant software suite (version 1.4.0.3, Max Planck Institute of Biochemistry, Department of Proteomics and Signal Transduction).

IV. SUPLEMENTAL INFORMATION

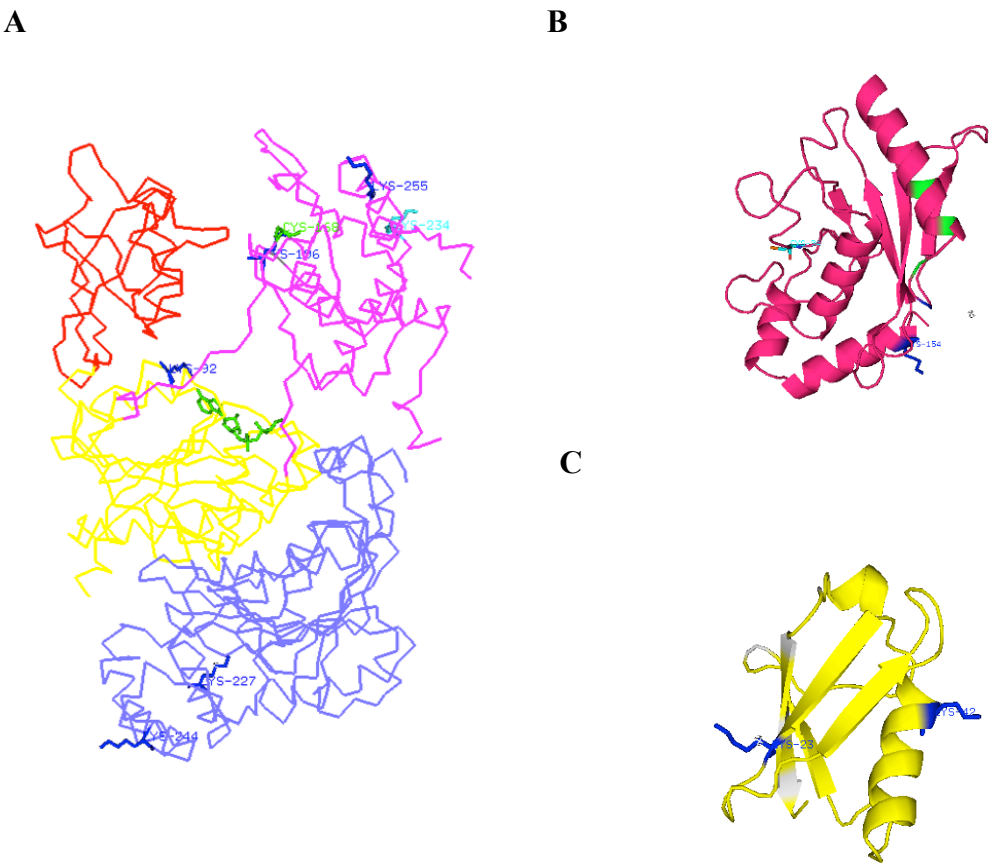


	Position	PEP	A1	A2	B1	B2	C1	C2	Sequence window of the Modified peptide
SAE2, SUMO activating enzyme subunit 2	525	1.51E-252	✓		✓				GSILTVEDLQQELSC**K**INVKHREEFDEEKEP
	92	7.31E-65	✓	✓	✓	✓			RFRPNININRSYHANV**K**NPEFDVDFFKQFDVV
	255	2.83E-57	✓	✓	✓	✓			FGSNIEAALSNEETW**K**NRRRPRPIYSKDVLP
	163	4.53E-49			✓	✓			KGKTECYECQTKPAP**K**TYPVCTITSTPTKFV
	619	1.97E-43	✓	✓	✓	✓			EDDDIMEVENPMMVS**K**KKIRVE_____
	595	1.49E-39	✓	✓	✓	✓			ASKKRRLSETEASNH**K**KETENVESEDDDIME
	234	3.15E-30	✓	✓	✓				VFERSEDEDIEQYGR**K**IYDHVFGSNIEAALS
	159	4.91E-24			✓	✓			TVHIKGKTECYECQT**K**PAPKTYPVCTITSTP
	196	3.92E-20	✓	✓	✓	✓			IVWAKDLLFAKLFGD**K**NQDNDLNVRSNNSAS
	417	6.00E-20	✓	✓					VDKFRMTYCLEHPSK**K**LLLMPIEPYEPNPAC
	214	2.43E-17			✓	✓			DNDLNVRSNNSASSS**K**ETEDVFERSEDEDIE
	176	3.30E-14			✓				APKTYPVCTITSTPT**K**FVHCIVWAKDLLFAK
	150	1.87E-10			✓				GTTGFLGQVTVHIKG**K**TECYECQTKPAPKTY
	416	4.44E-08			✓	✓			DVDKFRMTYCLEHPS**K**KLLLMPIEPYEPNPA
	14	5.78E-05			✓	✓			__MATQQQQSAIKGA**K**VLMVGAGGIGCELLK
	338	0.00045483			✓				FFAKRKKEIGHLTFD**K**DDQLAVEFVTAAANI
	191	0.00084444			✓	✓			KFVHCIVWAKDLLFA**K**LFGDKNQDNDLNVRS
	529	0.0343299			✓				TVEDLQQELSCKINV**K**HREEFDEEKEPEGMV
	148	0.041829			✓				ESGTTGFLGQVTVHI**K**GKTECYECQTKPAPK
	369	0.0532651		✓					RAESFGIPLHSLFEA**K**GIAGNIVHAVATTNA
SCE1	146	2.64E-80	✓		✓				QTDGYHLFCQDPVEY**K**KRVKLQSKQYPALV_
	154	4.88E-06	✓	✓					CQDPVEYKKRVKLQS**K**QYPALV_____
	28	0.00248594		✓					ERKSWRKNHPHGFVA**K**PETGQDGTVNLMVWH
SAE1a	244	1.35E-06			✓	✓			GECSLLDLARVLEIK**K**QLCEANSVSESHIPD
	227	2.94E-05	✓		✓	✓			FAMRVIEVFEESEGR**K**HGECSLLDLARVLEI
	98	6.02E-05			✓				VYSGKTVAEICSDSL**K**DFNPMVRVSVEKGDL
	143	0.0471088			✓				VIGYGSRATKKYVNE**K**CRKLKKRVAFYTVDC
SUMO	42	1.82E-19	✓	✓	✓	✓			GNEVFFRIKRSTQLK**K**LMNAYCDRQSVDMNS
	23	8.88E-13	✓	✓	✓	✓			DKKPGDGGAHINLKV**K**GQDGNEVFFRIKRST
	10	1.31E-12			✓	✓			_____MSANQEEDK**K**PGDGGAHINLKVKGQ

Table Supplementary 1. ***Putative SUMOylation sites identified on SUMO machinery components.*** Position of the lysine modified and PEP (posterior error probability) is showed for each enzyme. The modified lysine of the identified peptide is highlighted in black. In addition, it is shown in which samples the modified peptide has been identified: precipitate sample (A), SAE2 modififed purified (B) and SAE2 unmodified (C).



Supplemental figure S1. ***Mapped SUMOylation sites from SUMO E1 and SUMO E2***. Sequence alignment of the protein sequences of the SUMO activating enzyme (A) and SUMO conjugation enzyme (B) from *Arabidopsis thaliana*, *human* and *Saccharomyces cerevisiae* is shown. Black background and white letters corresponds to identical amino acids. Mapped SUMO attachment sites are indicated with red letters. Identified Lys-154 corresponds to the Lys-153 identified as a Smt3 attachment in ScUbc9 and is indicated with green dot.



Supplemental figure S2. ***Schematic localization of the SUMOylation sites identified.*** (A) Distribution of SAE2/SAE1 SUMOylation sites is shown. The SUMOylation sites located at the C-terminus are not shown due to the unresolved structure of SAE2 C-terminal tail. Distribution of UFD-Ct, Cys-Catalytic and adenylation domain, is represented in red, violet and yellow respectively. SAE1A is represented in blue. (B,C) SCE, SUMO conjugating enzyme and SUMO SUMOylation sites distribution among the structure is shown.

V. REFERENCES

Bencsath, K.P., Podgorski, M.S., Pagala, V.R., Slaughter, C.A., and Schulman, B.A. (2002). Identification of a multifunctional binding site on Ubc9p required for Smt3p conjugation. Journal of Biological Chemistry *277*, 47938–47945.

Bruderer, R., Tatham, M.H., Plechanovova, A., Matic, I., Garg, A.K., and Hay, R.T. (2011). Purification and identification of endogenous polySUMO conjugates. EMBO Reports *12*, 142–148.

Castano-Miquel, L., Segui, J., Manrique, S., Teixeira, I., Carretero-Paulet, L., Atencio, F., and Lois, L.M. (2013). Diversification of SUMO-Activating Enzyme in Arabidopsis: Implications in SUMO Conjugation. Molecular Plant *6*, 1646–1660.

Castaño Miquel, L., Seguí, J., and Lois, L.M. (2011). Distinctive properties of ArabidopsisSUMO paralogues support the in vivopredominant role of AtSUMO1/2 isoforms. Biochem. J. *436*, 581–590.

Colby, T., Matthaei, A., Boeckelmann, A., and Stuible, H.-P. (2006). SUMO-conjugating and SUMO-deconjugating enzymes from Arabidopsis. Plant Physiology *142*, 318–332.

Conti, E., Uy, M., Leighton, L., Blobel, G., and Kuriyan, J. (1998). Crystallographic analysis of the recognition of a nuclear localization signal by the nuclear import factor karyopherin alpha. Cell *94*, 193–204.

Elrouby, N., and Coupland, G. (2010). Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify Arabidopsis proteins implicated in diverse biological processes. Proc. Natl. Acad. Sci. U.S.a. *107*, 17415–17420.

Ho, C.-W., Chen, H.-T., and Hwang, J. (2011). UBC9 autosumoylation negatively regulates sumoylation of septins in Saccharomyces cerevisiae. J. Biol. Chem. *286*, 21826–21834.

Johnson, E.S., and Gupta, A.A. (2000). An E3-like Factor that Promotes SUMO Conjugation to the Yeast Septins. Cell *106*, 735–744.

Klug, H., Xaver, M., Chaugule, V.K., Koidl, S., Mittler, G., Klein, F., and Pichler, A. (2013). Ubc9 sumoylation controls SUMO chain formation and meiotic synapsis in Saccharomyces cerevisiae. Molecular Cell *50*, 625–636.

Knipscheer, P., Flotho, A., Klug, H., Olsen, J.V., van Dijk, W.J., Fish, A., Johnson, E.S., Mann, M., Sixma, T.K., and Pichler, A. (2008). Ubc9 sumoylation regulates SUMO target discrimination. Molecular Cell *31*, 371–382.

Knuesel, M., Cheung, H.T., Hamady, M., Barthel, K.K.B., and Liu, X. (2005). A method of mapping protein sumoylation sites by mass spectrometry using a modified small ubiquitin-like modifier 1 (SUMO-1) and a computational program. Molecular & Cellular Proteomics *4*, 1626–1636.

Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., and Yanagawa, H. (2008). Six Classes of Nuclear Localization Signals Specific to Different Binding Grooves of Importin. Journal of Biological Chemistry *284*, 478–485.

Liu, M., Shi, S., Zhang, S., Xu, P., Lai, J., Liu, Y., Yuan, D., Wang, Y., Du, J., and Yang, C. (2014). SUMO E3 ligase AtMMS21 is required for normal meiosis and gametophyte development in Arabidopsis. BMC Plant Biol. *14*, 153.

Lois, L.M., and Lima, C.D. (2005) . Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. Embo J. *24*, 439–451.

Miller, M.J., Scalf, M., Rytz, T.C., Hubler, S.L., Smith, L.M., and Vierstra, R.D. (2013). Quantitative Proteomics Reveals Factors Regulating RNA Biology as Dynamic Targets of Stress-induced SUMOylation in Arabidopsis. Molecular & Cellular Proteomics *12*, 449–463.

Miller, M., and Vierstra, R. (2010). Mass spectrometric identification of SUMO substrates provides insights into heat stress-induced SUMOylation in plants. Plant Signal Behav *6*, 130–133.

Miller, M.J., Barrett-Wilt, G.A., Hua, Z., and Vierstra, R.D. (2010). Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in Arabidopsis. Proc. Natl. Acad. Sci. U.S.a. *107*, 16512–16517.

Moutty, M.C., Sakin, V., and Melchior, F. (2011). Importin / mediates nuclear import of individual SUMO E1 subunits and of the holo-enzyme. Molecular Biology of the Cell *22*, 652–660.

Palancade, B., and Doye, V. (2007). Sumoylating and desumoylating enzymes at nuclear pores: underpinning their unexpected duties? Trends in Cell Biology *18*, 174–183.

Pichler, A., Knipscheer, P., Oberhofer, E., van Dijk, W.J., Körner, R., Olsen, J.V., Jentsch, S., Melchior, F., and Sixma, T.K. (2005). SUMO modification of the ubiquitin-conjugating enzyme E2-25K. Nat Struct Mol Biol *12*, 264–269.

Saracco, S.A., Miller, M.J., Kurepa, J., and Vierstra, R.D. (2007). Genetic Analysis of SUMOylation in Arabidopsis: Conjugation of SUMO1 and SUMO2 to Nuclear Proteins Is Essential. Plant Physiology *145*, 119–134.

Song, J., Tan, H., Perry, A.J., Akutsu, T., Webb, G.I., Whisstock, J.C., and Pike, R.N. (2012). PROSPER: an integrated feature-based tool for predicting protease substrate cleavage sites. PLoS ONE *7*, e50300.

Sriramachandran, A.M., and Dohmen, R.J. (2014). SUMO-targeted ubiquitin ligases. Biochimica Et Biophysica Acta (BBA)-

Tatham, M.H. (2001). Polymeric Chains of SUMO-2 and SUMO-3 Are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9. Journal of Biological Chemistry *276*, 35368–35374.

Tatham, M.H., Geoffroy, M.C., Shen, L., and Plechanovova, A. (2008). RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation : Abstract : Nature Cell Biology. Nature Cell

Truong, K., Lee, T.D., and Chen, Y. (2012a). Small Ubiquitin-like Modifier (SUMO) Modification of E1 Cys Domain Inhibits E1 Cys Domain Enzymatic Activity. Journal of Biological Chemistry *287*, 15154–15163.

Truong, K., Lee, T.D., Li, B., and Chen, Y. (2012b). Sumoylation of SAE2 C Terminus Regulates SAE Nuclear Localization. Journal of Biological Chemistry *287*, 42611–42619.

Ulrich, H.D. (2008). The Fast-Growing Business of SUMO Chains. Molecular Cell *32*, 301–305.

Vertegaal, A. (2011). Uncovering ubiquitin and ubiquitin-like signaling networks. Chemical Reviews.

Vertegaal, A.C.O., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M., and Lamond, A.I. (2006). Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. Molecular & Cellular Proteomics *5*, 2298–2310.

Wilkinson, K.A., and Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. Biochem. J. *428*, 133–145.

CAPÍTULO-VI

Identification of SUMO protein targets involved in the developmental transition from dormant to non-dormant seed

Identification of SUMO protein targets involved in the developmental transition from dormant to non-dormant seed

Laura Castaño-Miquel[1], Alfred C.O. Vertegaal[2], Maria Lois[1]

1 Center for Research in Agricultural Genomics CRAG (CSIC-IRTA-UAB-UB), Edifici CRAG-Campus UAB, Bellaterra (Cerdanyola del Vallés), 08193 Barcelona, Spain.
2 Department of Molecular Cell Biology, Leiden Univeristy Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

Resumen. La dormancia de las semillas es un característica intrínseca de las semillas que previene la germinación de las semillas en condiciones favorables de luz, humedad y temperatura. Desde el punto de vista agrícola, la dormancia presenta un problema cuando es demasiada débil o fuerte, produciendo grandes pérdidas económicas. Por esa razón, es crucial entender los mecanismos por los que está regulada con el fin de desarrollar una agricultura sostenible. Consecuentemente, se ha descrito el papel principal de dos hormonas el ácido abscísico y las giberelinas (ABA y GB), manteniendo o rompiendo la dormancia respectivamente, es la balanza entre ambas hormonas la que regula la dormancia de la semilla. Curiosamente, se ha descrito la interacción entre la vía de señalización de ABA y las modificaciones post-traduccionales de SUMO. Además, estudios en nuestro laboratorio han demostrado que la SUMOilación es un regulador negativo de la dormancia de las semillas.

En este contexto, en el presente estudio se ha estudiado el papel de SUMO durante la transición de un estado dormante a un estado no-dormante de una semilla. El análisis de la dinámica de SUMOilación ha mostrado un diferente patrón de SUMOilación en ambas condiciones. Se observa una mayor acumulación de conjugados de SUMO en semillas dormantes, que desaparecen rápidamente durante la transición a semillas no-dormantes.

Con el fin de entender el papel que juega SUMO en esta etapa hemos desarrollado un método eficiente para la purificación de sustratos de SUMO en semillas dormantes y no dormantes de plantas que sobre-expresan HIS_6-SUMO1. Las fracciones enriquecidas fueron obtenidas mediante una cromatografía de afinidad en condiciones desnaturalizantes y analizadas por HPLC-MS/MS. Se han identificado hasta 52 putativos sustratos de SUMO, solamente tres proteínas se han identificado en ambas condiciones. Se han identificado sustratos putativos de SUMO específicos para cada estado fisiológico analizado, 27 en semillas dormantes y 22 en semillas no-dormantes.

La lista de candidatos de SUMO se encuentra repartida en diferentes compartimentos celulares indicando que SUMO regula una amplia gama de procesos celulares. Este trabajo ha puesto de manifiesto, el papel de SUMO durante la transición de un estado dormante al estado no-dormante de una semilla, los resultados sugieren que existe una regulación específica por SUMO dependiente de la etapa desarrollo. Sin embargo, se necesitan más estudios para aclarar las consecuencias moleculares de la conjugación SUMO durante este proceso.

ABSTRACT

Seed dormancy is and adaptive trait that prevents the seed germination under optimal conditions and is regulated by complex genetic and environmental factors. From an agriculture perspective, weak or strong dormancy result in large productivity losses. For that reason, understanding the mechanisms affecting dormancy is crucial for developing a sustainable agriculture. Consequently, it is well known the major participation of two plant hormones in dormancy germination, ABA, and, release, GA through counterbalanced mechanims. Interestingly, it has been described the interplay between the ABA pathway and SUMO post-translational modifications. Furthermore, previous studies in our lab have elucidated that SUMOylation is a negative regulator of seed dormancy in *Arabidopsis*, consistently with its role as a negative regulator of ABA signalling.

In this context, in the present study we have further analysed the role of SUMO during the transition from dormant to non-dormant seeds. We have evaluated the SUMOylation dynamics in dormant and non-dormant seed and observed the presence of distinctive SUMO conjugation patterns. Dormant seeds are enriched in SUMO conjugates that disappear rapidly during dormancy release, evidencing the role of SUMO in the transition from dormant to non-dormant seeds.

To understand the SUMO function during this process, we have developed an efficient method to purify SUMO conjugates from dormant and non-dormant seeds. Taking advantage of plants overexpressing HIS_6-SUMO1, enriched and purified fractions from niquel affinity chromatography have been analysed by HPLC-MS/MS. A total of 52 putative SUMO targets have been identified and only three are present in both conditions. Stage specific putative SUMO substrates were found, 27 from dormant and 22 from non-dormant seeds. The identification of SUMO substrates candidates in different cellular compartments indicates a wide range of cellular processes that are modulated by SUMO. This is the first report that describes the role of SUMO during the transition from dormant to non-dormant seeds, elucidating that a specific SUMO regulation exists depending on the development stage. However, further studies are needed to clarify the molecular consequences SUMO conjugation during this process.

I. INTRODUCTION

The plant seed structure ensures the development and survival of the embryo guarantying the initiation of the next plant generation. The quality and proprieties of the seeds are essential traits for agriculture. Seed dormancy is an adaptive trait that prevents seed germination under favorable conditions such as optimal light, water and temperature. From an agronomical point of view, rapid and homogenous germination and seedling establishment is desirable, however some degree of dormancy is necessary in cereal crops to avoid germination of grains while still in the parent plant (pre-harvest sprouting), which results in major losses in the agriculture industry (Bewley, 1997). The release of seed dormancy is poorly understood and is regulated by complex interactions between environmental and genetic factors. There are two main key regulators ABA and GA hormones (abscisic acid and gibberellin), that act as antagonists. It is well known the role of the balance between ABA and GA, which determines whether the seed remains dormant or germinates (Koornneef et al., 2002). However, another factors have a role in dormancy release such endogenous signals, ethylene and nitric oxide, and post-translational modifications less known and documented (Arc et al., 2013).

Small-Ubiquitin-related MOdifier, SUMO, is a post-translational modifier belonging to the Ubiquitin family. SUMOylation consists of a covalent and reversible SUMO attachment to a lysine residue in the substrate, allowing the creation of "ON" and "OFF" states that regulates many biological processes very quickly in response to different biological conditions. A conserved enzymatic cascade, similar to the Ubiquitin system, carries out this conjugation with three crucial steps: activation, conjugation and ligation, involving the E1, E2 and E3 enzymes. Recently studies, have been described the interplay between ABA and SUMO at different levels. On one hand, mutants of *At*SIZ1 E3 ligase in response to drought conditions have 41% de-regulated of ABA related genes (Catalá et al., 2007). On the other hand, plants with impaired SUMOylation shows hypersensitivity to ABA (Lois, 2003). Consistently, the fundamental transcription factor of the ABA pathway ABI5 has been shown to be SUMOylated, ABI5 SUMOylation inactivates ABI5 and protects it in front degradation mediated by the proteasome (Miura et al., 2009). One of the major biological processes regulated by ABI5 is the promotion of growth arrest of the embryo during germination (Molina et al., 2002). On the other hand, nitric oxide, NO, accumulation promotes dormancy release by the activation of the ABA catabolism, reducing the ABA levels and inducing the germination of the embryo (Erwann Arc, 2013). Interestingly, the activity of a key enzyme in NO production, the nitrate reductase, has been shown to be stimulated by SUMOylation

(Park et al., 2011), suggesting that SUMOylation could control seed dormancy by regulating two signalling pathways, ABA and NO, in a coordinated manner. Unpublished results generates in our group have demonstrated that SUMOylation is a negative regulator of seed dormancy in Arabidopsis.

In this context, we aimed to investigate the role of SUMO during the transition from dormant to non-dormant seeds. Our main objective was the identification of SUMO substrates that would be key effectors in seed dormancy maintenance.

The identification of SUMO substrates is challenging due to the low abundance of SUMO targets and the high SUMO protease activity. To solve this problem, many researchers have proposed various approaches for SUMO conjugate purification; based on epitope-tagged SUMOs (such us HIS_6-tag, the Myc-tag, the HA-tag, FLAG-TEV…), using anti-SUMO antibodies or with SUMO affinity traps based on SIMs, SUMO interacting motifs, used to capture endogenous SUMOylated proteins (Vertegaal, 2011). Based on this, different SUMO proteomics approaches have identified hundreds of putative SUMO targets in mammalian, yeast and plants under different physiological conditions (Bruderer et al., 2011; Budhiraja et al., 2009; Denison et al., 2005; Galisson et al., 2011; Golebiowski et al., 2009; Miller et al., 2013; Miller and Vierstra, 2010; Schimmel et al., 2014; Vertegaal, 2004; Vertegaal et al., 2006).

Here, we show that the SUMO conjugate pattern observed in dormant dry seeds is replaced by a non-dormant seed specific pattern, suggesting a regulatory role of SUMO during seed dormancy and germination. In order to identify the proteins that are modified by SUMO under these two distinct physiological stages, we have developed an efficient method to purify SUMO conjugates from dormant and non-dormant seeds from *At*SUMO1 overexpressing plants and consisting in a stringent Ni^{2+} affinity purification combined with liquid chromatography and subsequent tandem mass spectrometry analysis. A total of 52 proteins have been identified, and only three are present in both physiological stages. Stages specific SUMO targets were found, 27 from dormant seeds and 22 from non-dormant seeds. To our knowledge, this is the first report showing an important role of SUMO conjugation in the transition from a dormant stage to an active growing stage.

II. RESULTS AND DISCUSSION

Dormant and non-dormant seeds display distinct SUMO conjugation patterns

In order to evaluate a potential role of SUMOylation in transition from dormant to non-dormant seed states, we analysed SUMO conjugation patterns of dormant and non-dormant seeds. Freshly harvested dry seeds were treated as dormant seeds. A fraction of the same seed batch was stratified (incubation at 4ºC in darkness in the presence of water for 48 hours) in order to break dormancy, non-dormant seeds. Protein extracts from dry and stratified seeds from Col-0 and plants overexpressing a N-terminal 6-His-tag *At*SUMO1 version under the constitutive Cauliflower Mosaic virus 35S promoter (CaMV 35S) was examined. Immunoblot analysis of protein SUMO levels was performed with polyclonal anti-SUMO1 antibodies (Figure 1 A). The results showed that the high molecular weight SUMO conjugates that are accumulated in dormant seeds are sensitive to stratification and diminished in non-dormant seeds. The high molecular weight SUMO conjugates could correspond to poly- or mono-SUMOyaltion of substrates of variable sizes. Other predominant proteins recognized by the anti-SUMO1 antibodies are the cruciferins that are the major storage proteins in Arabidopsis seeds, although we can not confirm the specificity of this immunoreactivity. Free SUMO1 is detected only in HIS-SUMO1 plants, consistently with a reduction in free SUMO accumulation detected during seed maturation. Although SUMO1 is overexpressed in HIS-SUMO1 plants, the excess of free-SUMO1 does not result in an increase of SUMO conjugates when compared to WT plants, suggesting that other components of the SUMOylation pathway are rate-limiting at this developmental stage.

Several molecular mechanisms could account for the reduction of SUMO conjugates upon stratification of dry seeds. One possibility might be that SUMO conjugates undergo deconjugation by the action of SUMO specific isopeptidases. Another possibility could involve protein degradation via proteosoma 26S mediated by initial SUMO conjugation. In order to address this question, dry seeds were stratified in presence or absence of cysteine protease inhibitors (N-ethylmaleimide, iodoacetamide and leupeptin) or the proteasome inhibitor MG132. Immunoblot analysis of SUMO conjugates levels did not reveal any significant differences between stratified seeds treated or not with protease or proteasome inhibitors (Figure 1B). Taking into account that stratification is a long treatment and the seed structure is organized to protect the embryo, we can not discard that poor penetration of the tested inhibitors in the seed tissues and/or loss of effectiveness during stratification could

account for these negative results. Further studies should address these possibilities in order to elucidate the molecular basis of SUMO conjugate dynamics during dormancy release.

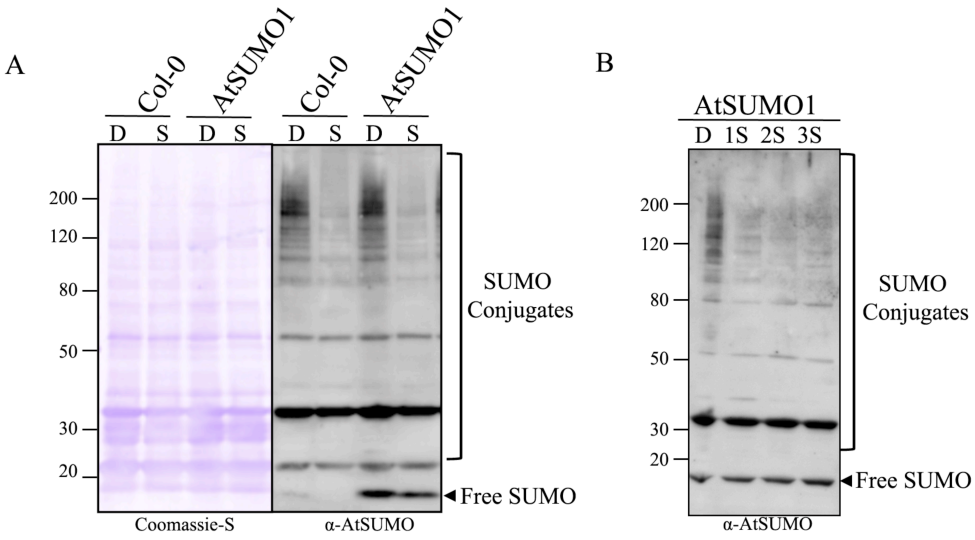


Figure 1. High molecular weight SUMO conjugates accumulated in dry seeds diminish during seed stratification. ***(A)*** Eighteen μg of total protein extracts from stratified (S) and dry (D) seeds from Col-0 and *At*SUMO1 plants were resolved by SDS-PAGE and examined by immunoblot analysis with anti-SUMO antibodies. Coomassie stained membrane was used to confirm equal protein loading. ***(B)*** *SUMO* protein patterns present in dry and stratified seeds from AtSUMO1 plants with different treatments are shown. D: Dry seeds; 1S: Stratified seeds; 2S: Stratified seeds in presence of protease inhibitors; 3S: Stratified seeds in presence of MG-132. Samples were subjected to SDS-PAGE and immunoblot analysis anti-SUMO1 antibodies.

Purification of SUMO conjugates

In order to elucidate the biological role of SUMO conjugation during the transition from dormant to non-dormant states, we aimed to identify the seed proteins that are conjugated by SUMO1 under these two physiological stages.

For this purpose, a stringent Ni^{2+}-affinity purification of SUMO conjugates from dry and stratified seeds from His_6-*At*SUMO1 overexpressing and Col0 plants was performed. To reduce the amount of proteins bound non-specifically to the column and to minimize deconjugation by endogenous SUMO proteases, protein extraction and purification was done in presence of strong denaturants (6M guanidine or 9M urea) and protease inhibitors. Equal amounts of protein extracts were used in each purification assay and affinity-purified proteins were collected in two elution fractions and subjected to immunoblot analysis with anti-SUMO1 antibodies. After purification, free SUMO and SUMO conjugates were significantly enriched in both elution fractions from *At*SUMO1 OE dry seeds, but not from Col0 seeds, indicating that the purification method was selective for recombinant His-SUMO1. Proteins purified from *At*SUMO1 OE stratified seeds displayed SUMO conjugates enrichment only in elution 1 but not in elution 2, consistently with a lower amount of SUMO conjugates in

stratified seeds (Figure 2, left panel). While the immunoblot showed a general enrichment of SUMOylated proteins from *At*SUMO1 OE samples, these proteins do not correspond to increased visible bands in the ponceau stained membrane (Figure 2, right panel), indicating the presence of some unspecific proteins bound to the column that were more abundant than SUMO conjugates.

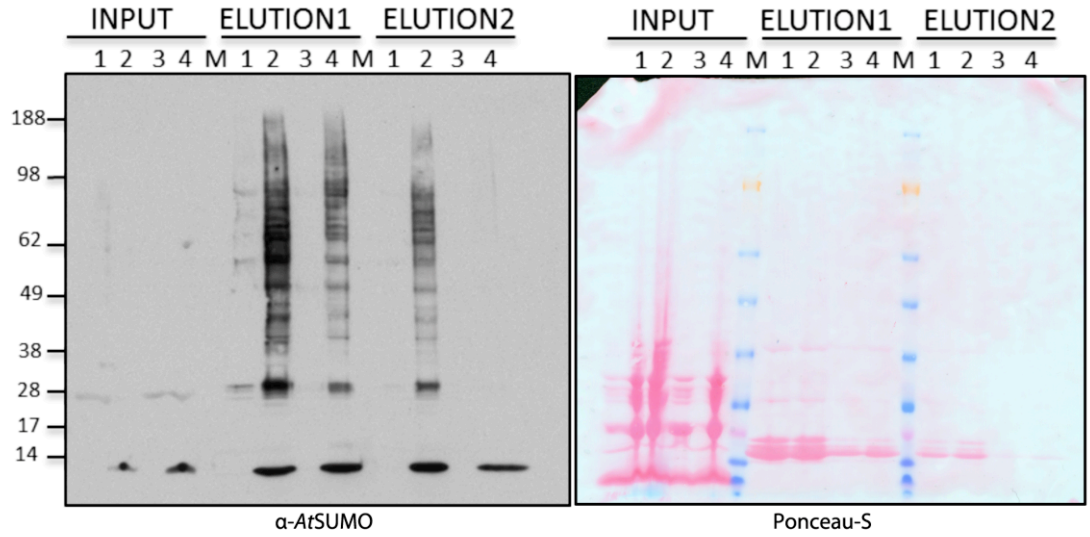


Figure 2: ***Affinity purification of SUMO1 conjugates from Arabidopsis seeds.*** Protein extracts from dry and stratified seeds from Col-0 and *At*SUMO1 OE plants were subjected to Ni-NTA affinity chromatography. Referred as 1: dry seeds from Col-0; 2: dry seeds from *At*SUMO1; 3: stratified seeds from Col-0; 4: stratified seeds from *At*SUMO1. The input and elution fractions were resolved by SDS-PAGE and subjected to immunoblot analysis with anti-SUMO1 antibodies (left panel). Ponceau stained membrane is showed (right panel).

Identification of SUMO1 conjugates by MS analysis

To identify endogenous SUMOylated proteins, affinity-purified fractions from Col-0 and *At*SUMO1 samples were concentrated and digested with trypsin and Lys-C. The resulting peptides were separated by HPLC and analysed by tandem mass spectrometry. Three biological replicates and two technical replicates were done. The resulting sequence peptides were searched against the Arabidopsis genome database. Potential SUMOylation substrates were selected based on the following criteria: 2 or more unique peptides were retrieved, detected in all three biological replicates from SUMO1 OE seeds and only in two biological replicates from Col0 seeds, and displaying at least 2-fold enrichment in SUMO1 OE samples with respect to Col0 samples. A total of 52 SUMO substrates candidates were selected distributed as follows: 27 putative SUMO targets in dormant seeds and 22 in non-dormant seeds (table S1). There was no significant overlap between the sets of putative SUMO conjugates, only 3 were found in both conditions as evidenced by Venn diagram analysis

(Figure 3A), suggesting that in addition to the overall reduction in SUMO conjugates during dormancy release, there is also a replacement in SUMO conjugates during this developmental transition.

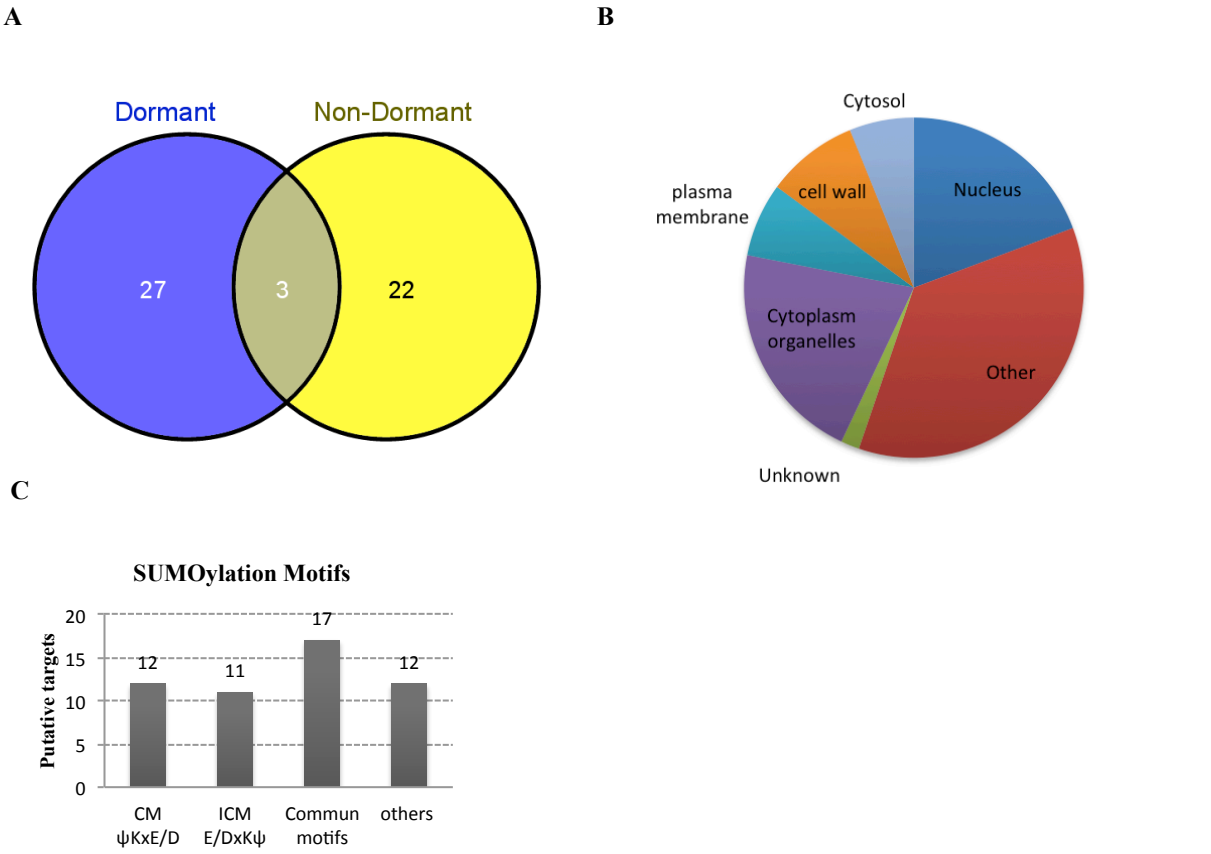


Figure 3. ***Annotation and characterisation of Arabidopsis SUMO substrates candidates by LC-MS/MS. (A)*** Venn diagram representation of putative SUMO substrates distributed between dormant and non-dormant seeds**.** 27 putative SUMO targets are specific to dry seeds, 22 SUMO targets candidates are specific to stratified seeds, and only 3 proteins are common in both physiological stages (SIZ1, GTE4 and DAA1). ***(B)*** Gene ontology enrichment analysis according to cellular component of the identified SUMO substrates was performed. ***(C)*** Distribution of the different types of SUMO attachment motifs over the putative targets identified. CM: consensus motif; ICM: inverted motif.

The set of putative SUMO substrates were found in the nucleus as well as the cytoplasm and organelles (Figure 3B) and cover a wide range of functions involved in metabolic processes, response to abiotic or biotic stimulus, protein metabolism and response to stress (Supplemental Figure S1). Supporting the recent evidence that SUMOylation is involved in functions in a wide range of cell compartments (Elrouby and Coupland, 2010). The analysis of the SUMO attachment sites, shows that 40 of the putative SUMO targets, a 76,9%, presents consensus and/or inverted SUMOylation motifs (Figure 3C), indicating that there is an enrichment of the proteins containing SUMOylation motifs over the SUMO candidates identified.

The highest-ranking candidates correspond to six proteins present exclusively in samples from *At*SUMO1 OE seeds and that provided the identification of at least four unique peptides (table I). The presence of consensus SUMOylation sites was analysed using SUMOplot software and resulted in the identification of conserved SUMO attachment sites in all six proteins. In addition, our proteomic approach is supported by the presence of previously characterized SUMO substrates, SIZ1, GTE4 and GTE3, among these high confidence six candidates (Miller et al., 2013; 2010). In particular, SIZ1 was shown to be transiently SUMOylated upon heat stress as a rapid response, although the molecular consequences of SIZ1 SUMOylation regarding its E3 ligase activity remains unknown (Miller and Vierstra, 2010). In seeds, we have found that SIZ1 is steadily SUMOylated from dormant to non-dormant seed states, although the relative enrichment is 3.6-fold higher in dormant seeds, suggesting that SIZ1 deSUMOylation may be necessary for dormancy release.

Table I. ***TOP 6 of SUMO targets candidates***

Protein Name	*AGI code*	*D*	*S*	*Seq. Coverage %*	*Description*	*Subcellular localization*	*SUMOylation consensus motif*	*Ratio Dry vs Stratified*
GTE4*	AT1G06230	✓	✓	23.9	Global transcription factor, encodes a bromodomain-containing protein.	Unknown cellular components	Lys^{193},Lys^{700}	4.1
SIZ1*	AT5G60410	✓	✓	20.1	DNA-binding protein with MIZ/SP-RING zinc finger, PHD-finger and SAP domain. SUMO pathway protein. E3 SUMO ligase.	Nucleus	Lys^{574},$Lys^{23,}$ Lys^{517}	3.6
DAA1 DUO1-Activated-ATPASE1	AT1G64110	✓	✓	23.2	Target promoter of the male germline-specific transcription factor DUO1. ATP binding	Plasma membrane	Lys^{757},Lys^{624},Lys^{547}, Lys^{541},Lys^{571}	2.9
GTE3*	AT1G73150	✓	-	13	Bromodomain and extra terminal domain family protein. Binds to acetyl-histone H3	Nucleus	Lys^{187}	
alpha/beta-Hydrolases	AT5G53050	-	✓	17.7	hydrolase activity, catalytic activity	Cytoplasm	Lys^{10}	-
APG7	AT5G45900	✓	-	12.5	Component of autophagy conjugation pathway. Required for proper senescence	Cytoplasm	Lys^{259},Lys^{133},Lys^{37}	-

* Proteins found by other proteomics analysis *in vivo* (Miller et al., 2013; 2010).
- Dry and stratified conditions are referred as D and S, respectively. The relative ratio between DRY/STRAT of the 3 common proteins is shown. Fold change was calculated by the average of normalized unique peptides of the enriched SUMO biological replicate in dry seeds compared to stratified seeds.
- SUMOplot were used to predict potential SUMO attachment sites. Only motifs with high probability and high score (>0.8) are shown (http://www.abgent.com/sumoplot).

Our proteomic analysis identified GTE3 and GTE4 global transcription factors. GTE4 have been identified in both conditions while GTE3 is only present in dormant conditions. The

GTE family has been poorly investigated. GTE1, 3, 4 and 7 have also been identified as a SUMO substrate by proteomic approaches in *Arabidopsis thaliana* seedlings subjected to heat and oxidative stress (Miller et al., 2013; 2010). GTE3 binds to acetyl-histone H3, although binding is reduced when GTE3 is SUMOylated (Garcia-Dominguez et al., 2008), supporting the role of SUMOylation in chromatin regulation in a spatial-temporal manner. The SUMOylation of DAA-1 has also been found in both seed states and, regardless it has shown to regulate a battery of cell-expressed genes required for the male gametogenesis (Borg et al., 2011), in silico expression analysis show that DAA1 is highly expressed in dry seeds. Similarly to SIZ1, GTE3 and DAA1 relative enrichment is higher in dry than in stratified seeds. Future experiments should analyse whether this is the result of a reduction in total protein, SUMO modified and unmodified, or if there is a specific reduction of the SUMOylated form. Considering the dramatic changes undergoing during this developmental transition, both scenarios are possible. These possibilities could also account for the state-specificity of the identified substrates.

Overall, further studies are required to test all the set of putative SUMO targets and understand the molecular consequences of their SUMOylation and the implications in the transition from dormant to non-dormant states. Our SUMO conjugate enrichment procedure described can be easily adapted to find SUMO substrates in other conditions or for other ubiquitin-like protein modifiers. Nevertheless, we envision that the addition of purification steps and/or the increment of stringency during the purification could contribute to reduce the unspecific bounding to the Ni^{2+} agarose as illustrated by the small differences observed in the Ponceau-stained proteins from SUMO1 OE and Col0 seeds. For example, to overcome this problem, the addition of purification step based on anti-SUMO antibodies affinity purification or the use of a different tag in addition of the HIS_6-tag could result in an optimized protocol. Furthermore, to favour the conjugation of the tagged SUMO instead of native SUMO, double homozygous *sum1-1 sum2-1* complemented with HIS6-SUMO1 might be used.

In conclusion, our results point to a major role of SUMO in the regulation of the transition from dormant to non-dormant seeds not reported before. Distinct SUMOylation patterns have been shown in both stages. We have identified a total of 52 putative SUMO substrates in dry and stratified seeds. The identification of putative SUMO substrates outside the nucleus indicates a wide range of cellular processes modulated by SUMO. Taking into account the low overlapping of SUMO conjugates between dormant and non-dormant stages, it appears that a specific SUMO regulation exists depending on the developmental stage. In addition,

the identification of previously validated SUMO substrates supports the other identified proteins as potential SUMO substrates. Further research on the putative SUMO substrates identified would be necessary to extend our knowledge of the role of SUMO during the transition form dormant to non-dormant seeds.

III. MATERIALS AND METODS

Seed harvesting and stratified treatments

Freshly harvested seeds were collected from 6-8 week old of Col-0 and *At*SUMO1 OE plants. Seeds were stratified in presence of water at 4ºC for 48 hours in the dark. For protein levels analysis around 50μL of seed volume was used. For SUMO1 conjugates purification 0,4 grams of Arabidopsis seeds were used for each replicate and condition. When indicated, stratification was performed in presence of 10μM of MG-132 or protease inhibitors (20mM of NEM, 10mM iodoacetamide and 1μg/μL of Leupetin).

Protein extraction and immunoblot

Total protein extracts were prepared from dry and stratified seeds. For analytical experiments, seeds were grinded using mortar and pestle in liquid nitrogen, total proteins were extracted at 4ºC for 1 hour in the urea lysis buffer containing 100mM Tris-HCl ph7.5, 8M Urea, 0.2% Triton X-100, 0,2% Sarkosyl, 1mM PMSF, 1μg/mL Pepstatin, 1μg/mL Leupeptin, 2mM NEM and 10mM Iodoacetamide. For SUMO conjugates purification, seeds were grinded and proteins were extracted at 4ºC for 1 hour in the guanidine lysis buffer containing 6M guanidine-HCl, 0.1M Na_2HPO_4/NaH_2PO_4 pH 8.0, 0.01M Tris-HCl pH 8.0, 0,2% Sarkosyl, 10mM Iodoacetamide, 0.2% Triton X-100 and cocktail inhibitors (Roche, cOmplete EDTA free). The crude extract was clarified by centrifugation at 11,000 rpm for 30 min at 4ºC. For immunoblot analysis protein samples were resolved on Novex 4-12% BisTris gradient gels using MOPS Buffer (Invitrogen) and subsequently transferred onto PDVF membranes (Milipore) and analysed by polyclonal anti-SUMO1 antibodies. Chemiluminescence signal from ECL plus reagent (GE Healthcare) was acquired with the LAS-4000 imaging system (Fujifilm).

Purification of SUMOylated Proteins

0.4 grams of stratified and dry seeds from Col-0 and *At*SUMO1 plants were extracted as described before. Protein extracts were passed through 0.45μm filter column and prepared for the purification step by adding 10mM Imidazole pH8.0, 100mM NaCL and 5mM β-Mercaptoethanol; equal amounts of protein were incubated with 400μL of Ni-NTA agarose at room temperature for 3 hours. After incubation the beads were washed four times with different buffers:

- Wash buffer 1: 6 M guanidine-HCl, 0.1M Na_2HPO_4/NaH_2PO_4 pH 8.0, 0.01M Tris-HCl pH 8.0, 5 mM β-mercaptoethanol and 0.1% Triton X-100.

- Wash buffer 2: 8 M urea, 0.1M Na_2HPO_4/NaH_2PO_4 pH 8.0, 0.01M Tris-HCl pH 8.0, 10mM imidazole pH 8.0, 5mM β-mercaptoethanol, 0.1% Triton X-100.
- Wash buffer 3: 8 M urea, 0.1M Na_2HPO_4/NaH_2PO_4 pH 6.3, 0.01M Tris-HCl pH 6.3, 10mM imidazole pH 7.0, 5mM β-mercaptoethanol.
- Wash buffer 4: 8 M urea, 0.1M Na_2HPO_4/NaH_2PO_4 pH 6.3, 0.01M Tris-HCl pH 6.3, and 5mM β-mercaptoethanol.

Bound proteins were eluted two times for 30 minutes in one bead volume of the elution buffer containing 7 M urea, 0.1M Na_2HPO_4/NaH_2PO_4, 0.01M Tris-HCl, pH 7.0, 250 mM imidazole pH 7.0.

Sample preparation for Mass spectrometry analysis

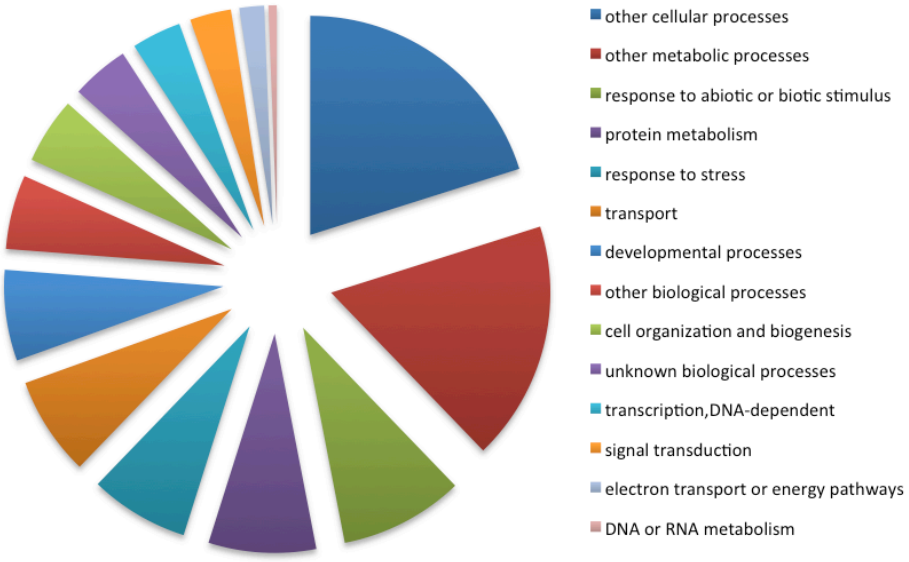
Eluted fractions from dry and stratified seeds from *At*SUMO1 OE and Col-0 plants were concentrated at room temperature using 50kDa cut-off filters until 25μL. 1M of ammonium bicarbonate (ABC) was added to the protein mixtures to reach 50mM ABC final concentration. Samples were reduced with 1mM dithiothreitol (DTT) and incubated 30 minutes at room temperature. Reaction mixtures were alkylated with chloroacetamide (CAA) to 5mM final concentration, and incubated 30 minutes at room temperature. Additional DTT was added to yield 6mM DTT, and incubated for 30 minutes at room temperature. Samples were subjected to endoproteinase Lys-C digestion for 4 hours at room temperature, in a 1:50 enzyme-to-protein ratio. Lys-C digested samples were diluted 4-fold by addition of 3 volumes of 50mM ABC and incubated overnight at room temperature with trypsin, 1:50 enzyme-to-protein ratio. The resulting peptide mixtures were acidified with 2% trifluoroacteic (TFA) and desalted and concentrated by C_{18} solid-phase extraction using 40μL of 80% acetonitrile and 0,1% TFA for elution. Vaccum dried samples were resuspended in 0,1% TFA and subjected to LC/MS using LTQ-Orbitrap (Thermo Fisher Scientific). Mass spectrometry files were processed with the MaxQuant software package (Cox et al., 2009).

IV. SUPLEMENTAL INFORMATION

Table Supplementary 1. ***List of putative SUMO Substrates identified in dormant and non-dormant seeds.***

	GENE	Localitzation	Protein name
Commun	AT1G06230		GTE4; transcription factor GTE4
	AT5G60410	nucleus	SIZ1, E3 SUMO-protein ligase SIZ1
	AT1G64110	plasma membrane	DAA1, Transcription factor DUO1
Dormant	AT1G73150	nucleus	GTE3; Transcription factor GTE3
	AT5G45900	cytoplasm, cytosol	ATG7; Ubiquitin-like modifier-activating enzyme atg7
	AT1G58230	mithocondria	WD40 and Beach domain-containing protein
	AT1G65930	plasma membrane, cytosol, chloroplast	cICDH; cytosolic NAD+- dependent isocitrate dehydrogenase
	AT3G24670	extracelular region	Pectin lyase-like superfamily protein
	AT4G18950	plasma membrane, cytosol	Integrin-linked protein kinase family
	AT1G58060	chloroplast	RNA helicase family protein
	AT3G06400	nucleus, cytosol	CHR11; CHROMATIN-REMODELING PROTEIN 11
	ATCG00120	chloroplast	ATP SYNTHASE SUBUNIT ALPHA
	AT2G05120	chloroplast	Nup133/Nup155-like;Nucleoporin
	AT1G08420	plasma membrane	BSL2; BRI1 supress OR1 (BSU1)-like 2
	AT3G54400	cell wall, extracellular region, apoplast	aspartyl protease family protein
	AT5G18640	extracelular region	alpha/beta-Hydrolases superfamily protein
	AT1G14850	nucleus, chloroplast, plasmodesma	NUP155; NUCLEOPORIN 155
	AT1G52160	nucleus	TRZ3; tRNAse Z3
	AT5G42270	chloroplast	VAR1, FTSH5; FtsH extracellular protease family
	AT1G72960	nucleus	RHD3; root hair defective 3 GTP-binding protein
	AT3G01290	mitochondria, vacuole, plasma membrane	HIR2; Hypersensitive induced reaction 2
	AT1G77000	nucleus	AtSPK2; Arabidopsis Homolog of human SKP2
	AT2G28470	extracellular region, cell wall	BGAL8; Beta-galactosidase 8
	AT2G41770	chloroplast, glogi aparatus	DUF288; protein of unknown function
	AT5G53850	cytoplasm, chloroplast	DEP1; dehdratase-enolase-phosphate-complex 1
	AT5G10290	plasma membrane	leucine-rich repeat transmembrane protein kinase family protein
	AT3G12740	golgi aparatus	ALIS1; Ala-interacting subunit 1
	AT1G74320	cytoplasm	Choline kinase
	AT2G41560	vacuole	ACA4; autoinhibited CA(2+)-ATAPase
	AT2G02070	nucleus, intracellular	AtIDD5, IDD5; Indeterminated(ID)-Domain 5
Non-dormant	AT5G37350	cytoplasm	Serine/threonine-protein kinase Rio1
	AT5G53050	cytoplasm	alpha/beta-Hydrolases superfamily protein
	AT5G56590	plasma membrane, cell wall	O-Glycosyl hydrolases family 17 protein
	AT3G13020	nucleus	hAT transposon superfamily protein
	AT4G11670	nucleus, mitochondria	DUF810; protein of unknown function
	AT2G35920	cytoplasm	RNA helicase family protein
	AT3G05380	nucleus	ALY2; Always early 2
	AT3G53870	cytoplasm, ribosome, intracellular, plasmodesm	Ribosomal protein S3 family
	AT3G08020	nucleus	PHD finger family protein
	AT5G04550	nucleus	DUFF668; Protein of unknown function
	AT4G24840	membrane, vacuole, Golgi aparatus	unknown molecular function
	AT5G43130	nucleus	TAF4; TBP-associated factor 4
	AT3G15940	golgi aparatus	UDP-Glycosyltransferase superfamily protein
	AT5G26751	cytosol, nucleus	SK11; Shaggy-related protein kinase alpha
	AT5G63840	extracellular region, chloroplast	RSW3; radial swelling 3
	AT2G03450	extracellular region	PAP9; purple acid phosphatase 9
	AT1G54385	unknown	ARM repeat superfamily protein
	AT1G09920	nucleus	TRAF-type finger related
	AT1G17690	nucleus	NOF1; nucleolar factor 1
	AT1G68910	nucleus	WIT2; WPP domain-interacting protein 2
	AT2G32700	nucleus	MUM1; mucilage-modified 1
	AT4G31880	nucleus, chloroplast	Tduor/PWWP/MBT superfamily protein

Supplemental Figure S1. ***Biological Functions of putative SUMO substrates in dry and stratified conditions.*** Distribution of putative SUMOylated proteins by enriched gene ontology (GO) in terms of biological functionality.



V. REFERENCES

Arc, E., Sechet, J., Corbineau, F., Rajjou, L., and Marion-Poll, A. (2013). ABA crosstalk with ethylene and nitric oxide in seed dormancy and germination. Front Plant Sci *4*, –63.

Bewley, J.D. (1997). Seed Germination and Dormancy. The Plant Cell Online *9*, 1055–1066.

B**org, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M., and Twell, D.** (2011). The R2R3 MYB Transcription Factor DUO1 Activates a Male Germline-Specific Regulon Essential for Sperm Cell Differentiation in Arabidopsis. The Plant Cell.

Bruderer, R., Tatham, M.H., Plechanovova, A., Matic, I., Garg, A.K., and Hay, R.T. (2011). Purification and identification of endogenous polySUMO conjugates. EMBO Reports *12*, 142–148.

Budhiraja, R., Hermkes, R., Muller, S., Schmidt, J., Colby, T., Panigrahi, K., Coupland, G., and Bachmair, A. (2009). Substrates Related to Chromatin and to RNA-Dependent Processes Are Modified by Arabidopsis SUMO Isoforms That Differ in a Conserved Residue with Influence on Desumoylation. Plant Physiology *149*, 1529–1540.

Catalá, R., Ouyang, J., Abreu, I.A., Hu, Y., Seo, H., Zhang, X., and Chua, N.-H. (2007). The Arabidopsis E3 SUMO ligase SIZ1 regulates plant growth and drought responses. Plant Cell *19*, 2952–2966.

C**ox, J., Matic, I., Hilger, M., Nagaraj, N., Selbach, M., Olsen, J.V., and Mann, M.** (2009). A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. Nat Protoc *4*, 698–705.

Denison, C., Rudner, A.D., Gerber, S.A., Bakalarski, C.E., Moazed, D., and Gygi, S.P. (2005). A proteomic strategy for gaining insights into protein sumoylation in yeast. Molecular & Cellular Proteomics *4*, 246–254.

Elrouby, N., and Coupland, G. (2010). Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify Arabidopsis proteins implicated in diverse biological processes. Proc. Natl. Acad. Sci. U.S.a. *107*, 17415–17420.

Erwann Arc, M.G.B.G.G.C.L.R. (2013). Nitric oxide implication in the control of seed dormancy and germination. Front Plant Sci *4*.

Galisson, F., Mahrouche, L., Courcelles, M., Bonneil, E., Meloche, S., Chelbi-Alix, M.K., and Thibault, P. (2011). A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. Molecular & Cellular Proteomics *10*, M110–M004796.

Garcia-Dominguez, M., March-Diaz, R., and Reyes, J.C. (2008). The PHD Domain of Plant PIAS Proteins Mediates Sumoylation of Bromodomain GTE Proteins. Journal of Biological ….

Golebiowski, F., Matic, I., Tatham, M.H., Cole, C., Yin, Y., Nakamura, A., Cox, J., Barton, G.J., Mann, M., and Hay, R.T. (2009). System-wide changes to SUMO modifications in response to heat shock. Sci Signal *2*, ra24.

Koornneef, M., Bentsink, L., and Hilhorst, H. (2002). Seed dormancy and germination. Current Opinion in Plant Biology *5*, 33–36.

Liu, M., Shi, S., Zhang, S., Xu, P., Lai, J., Liu, Y., Yuan, D., Wang, Y., Du, J., and Yang, C. (2014). SUMO E3 ligase AtMMS21 is required for normal meiosis and gametophyte development in Arabidopsis. BMC Plant Biol. *14*, 153.

Lois, L.M. (2003). Small Ubiquitin-Like Modifier Modulates Abscisic Acid Signaling in Arabidopsis. The Plant Cell *15*, 1347–1359.

Miller, M.J., Scalf, M., Rytz, T.C., Hubler, S.L., Smith, L.M., and Vierstra, R.D. (2013). Quantitative Proteomics Reveals Factors Regulating RNA Biology as Dynamic Targets of Stress-induced SUMOylation in

Arabidopsis. Molecular & Cellular Proteomics *12*, 449–463.

Miller, M., and Vierstra, R. (2010). Mass spectrometric identification of SUMO substrates provides insights into heat stress-induced SUMOylation in plants. Plant Signal Behav *6*, 130–133.

Miller, M.J., Barrett-Wilt, G.A., Hua, Z., and Vierstra, R.D. (2010). Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in Arabidopsis. Proc. Natl. Acad. Sci. U.S.a. *107*, 16512–16517.

Miura, K., Lee, J., Jin, J.B., Yoo, C.Y., Miura, T., and Hasegawa, P.M. (2009). Sumoylation of ABI5 by the Arabidopsis SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. Proc. Natl. Acad. Sci. U.S.a. *106*, 5418–5423.

Molina, L.L., Mongrand, S., and McLachlin, D.T. (2002). ABI5 acts downstream of ABI3 to execute an ABA-dependent growth arrest during germination - Lopez-Molina - 2002 - The Plant Journal - Wiley Online Library. The Plant ….

Park, B.S., Song, J.T., and Seo, H.S. (2011). Arabidopsis nitrate reductase activity is stimulated by the E3 SUMO ligase AtSIZ1. Nat Commun *2*, 400–.

Schimmel, J., Eifler, K., Sigurðsson, J.O., Cuijpers, S.A.G., Hendriks, I.A., Vries, M.V.-D., Kelstrup, C.D., Francavilla, C., Medema, R.H., Olsen, J.V., et al. (2014). Uncovering SUMOylation Dynamicsduring Cell-Cycle Progression Reveals FoxM1 as a Key Mitotic SUMO Target Protein. Molecular Cell *53*, 1053–1066.

Vertegaal, A.C.O. (2004). A Proteomic Study of SUMO-2 Target Proteins. Journal of Biological Chemistry *279*, 33791–33798.

Vertegaal, A. (2011). Uncovering ubiquitin and ubiquitin-like signaling networks. Chemical Reviews.

V**ertegaal, A.C.O., Andersen, J.S., Ogg, S.C., Hay, R.T., Mann, M., and Lamond, A.I.** (2006). Distinct and overlapping sets of SUMO-1 and SUMO-2 target proteins revealed by quantitative proteomics. Molecular & Cellular Proteomics *5*, 2298–2310.

RESUMEN GLOBAL DE RESULTADOS

Este trabajo se ha centrado en conocer los determinantes moleculares implicados en la conjugación de SUMO en plantas, desarrollar una estrategia para inhibir la conjugación de SUMO *in vivo,* y elucidar por primera vez la regulación por SUMO de la respuesta de defensa de las plantas frente patógenos necrótrofos y de la transición de semilla quiescente a germinativa. Con este fin se ha estudiado la SUMOilación en *A.thaliana* desde diferentes puntos de vista:

❖ El estudio molecular de los parálogos de SUMO en *Arabidopsis thaliana* ha demostrado que son funcionalmente más divergentes que sus ortólogos humanos:

- ✓ Entre los parálogos de SUMO no se da una conservación de la superficie de interacción con la enzima E1 activadora y E2 conjugadora de SUMO; a diferencia de lo que sucede en humanos donde se encuentra altamente conservada.
- ✓ Cada una de las isoformas de SUMO humanas interacciona no covalentemente con la enzima E2 conjugadora de SUMO (Capili and Lima 2007). En cambio, tan sólo *At*SUMO1 y 2 interaccionan no covalentemente con la enzima conjugadora SCE1, mientras que *At*SUMO3 y 5 no presentan esta capacidad.
- ✓ El aspártico 63 de *At*SUMO1 es esencial para la interacción no covalente con la enzima E2 conjugadora, pero no es suficiente.
- ✓ Al igual que en humanos, la interacción no covalente con *At*SCE1 está relacionada con la eficiencia de formación de cadenas de SUMO, pero no con la tasa de conjugación.
- ✓ Los ensayos *in vitro* han demostrado que *At*SUMO1 y 2 son las formas que se conjugan más eficientemente, seguidas de *At*SUMO3. No hemos detectado conjugación de *At*SUMO5 *in vitro*.
- ✓ La interacción no covalente entre la enzima activadora y SUMO está relacionada con la eficiencia de conjugación. La sustitución de los aminoácidos conservados en *At*SUMO1 por los residuos no conservados de las isoformas *At*SUMO3 y 5, tienen un efecto negativo en la tasa de conjugación, sugiriendo que la menor conjugación de estas isoformas es debida a un ineficiente reconocimiento de estas enzimas por parte de la enzima activadora.

❖ Control de la SUMOilación *in vivo* por la diversificación de la enzima activadora en *Arabidopsis:*

- ✓ La subunidad pequeña de la enzima activadora, SAE1, esta codificada por tres genes: *SAE1a*, *SAE1b1* y *SAE1b2*. Estos genes son resultados de dos duplicaciones independientes. Sin embargo, tan sólo existen dos isoformas de la enzima activadora de SUMO, E1a y E1b, debido a que los genes *SAE1b1* y *SAE1b2* corresponden a dos copias exactas duplicadas en tándem en el genoma de *Arabidopsis thaliana*.
- ✓ SAE1a y SAE1b están altamente conservadas entre ellas, tan sólo presentan un 20% de divergencia de secuencia. La secuencia más conservada corresponde a la región de interacción con la subunidad grande de la enzima activadora de SUMO.
- ✓ Mediante ensayos *in vitro* se ha determinado una mayor tasa de conjugación de la isoforma E1a que la E1b a altas temperaturas. Mientras que no se observan diferencias en la capacidad de formación del enlace tioéster con SUMO, sugiriendo que pasos posteriores a la activación deben afectar a la diferente tasa de conjugación observada.
- ✓ La localización de ambas isoformas es nuclear y dependiente de una señal de localización nuclear canónica, NLS1, situada en el extremo C-terminal de SAE2.
- ✓ Los mutantes *atsae1a* no muestran fenotipo en condiciones basales. Aunque, cuando estas plantas son sometidas a estrés térmico o de sequía se observa una menor acumulación de conjugados que en las plantas wilde-type.

❖ Inhibición de la SUMOilación mediante la expresión del dominio $SAE2^{UFDCt}$:

- ✓ El dominio $SAE2^{UFDCt}$ es esencial para la actividad de la enzima activadora de SUMO.
- ✓ La interacción no covalente entre $SAE2^{UFDCt}$ y la enzima conjugadora de SUMO, SCE1, da lugar a la inhibición de la conjugación de SUMO *in vitro* mediante la supresión de la interacción entre las enzimas E1 y E2 de SUMO.
- ✓ El dominio $SAE2^{UFDCt}$ localiza en el núcleo mientras que SCE1 está presente en núcleo y citoplasma. La co-expressión de ambas resulta en una localización nuclear de SCE, demostrando la interacción de ambas proteínas *in vivo.*

- ✓ La región LHEB2 de SAE2 está altamente conservada entre los ortólogos de plantas y está implicada en el reconocimiento de la enzima conjugadora de SUMO.
- ✓ La sobreexpresión del dominio $SAE2^{UFDCt}$ en *A.thaliana* da lugar a plantas con un fenotipo enano y floración temprana. Además estas plantas presentan unos niveles de conjugados de SUMO inferiores y una mayor expresión de SCE1 en comparación con plantas wild-type. Estas alteraciones son características de plantas que presentan la SUMOilación afectada, cómo el caso del mutante *siz1-3*. Validando la supresión de la interacción entre E1 y E2 de SUMO mediante la sobreexpresión del dominio $SAE2^{UFDCt}$ como una estrategia eficiente para inhibir la conjugación de SUMO *in vivo*.

❖ Es necesario un correcto funcionamiento de la maquinaria de SUMO para activar la respuesta frente a patógenos necrótrofos. Plantas con la SUMOilación disminuida muestran una mayor sensibilidad frente a *B.cinerea* y *P. cucumerina.*

❖ Regulación de la enzima activadora de SUMO mediante modificaciones post-traduccionales:

- ✓ Ocho aminoácidos del extremo C-terminal de SAE2 son procesados durante el desarrollo de la semilla.
- ✓ Los ocho aminoácidos procesados constituyen una señal de localización nuclear no canónica, NLS2, esencial para una completa localización nuclear de la enzima. Sugiriendo que está modificación puede ser un mecanismo de la regulación de la conjugación de SUMO *in vivo*, mediante la relocalización de la enzima activadora en condiciones específicas con el fin de favorecer la modificación específica de sustratos citoplasmáticos de SUMO.
- ✓ Se da una conservación del dominio C-terminal de SAE2 entre los ortólogos de SUMO en plantas.
- ✓ Se han identificado las lisinas modificadas por SUMO de diferentes componentes de la maquinaria de SUMO (SAE2, SAE1A, SCE1 y SUMO). Modificaciones similares han sido identificadas en el sistema de SUMOilación de levadura y mamíferos, sugiriendo que estas modificaciones podrían constituir un mecanismo conservado de regulación de la actividad, estabilidad y localización de estas enzimas *in vivo.*

❖ Diferente dinámica de conjugados de SUMO durante la ruptura de la dormancia:

- ✓ Existen diferencias en el patrón de conjugados de SUMO entre semilla dormante y no dormante. Durante la ruptura de la dormancia se observa una disminución significativa de conjugados de SUMO.
- ✓ Se ha desarrollado un método efectivo para la purificación de sustratos de SUMO en semilla dormante y no dormante, mediante condiciones desnaturalizantes y astringentes de plantas sobreexpresoras de HIS_6-SUMO seguido de LC-MS/MS.
- ✓ Se han identificación un total de 52 candidatos sustratos de SUMO. 27 y 22 sustratos potenciales de SUMO son específicos de semilla seca y estratificada, respectivamente. Tan sólo 3 son comunes en ambas condiciones, sugiriendo una modificación específica por SUMO en condiciones dormantes y no dormantes.

DISCUSIÓN GENERAL

El objetivo general de la presente tesis doctoral ha sido el estudio del sistema de SUMOilación en plantas desde las perspectivas molecular y biológica. Para ello se ha caracterizado bioquímica y funcionalmente los parálogos de SUMO y la enzima activadora de SUMO en *A.thaliana*. Los resultados presentan nuevos avances en el conocimiento sobre la actividad y la regulación de los componentes de la maquinaria de SUMO de *Arabidopsis*. A partir de los resultados obtenidos, se ha desarrollado una herramienta molecular que ha permitido descubrir nuevos procesos biológicos que son regulados por la SUMOilación.

Caracterización funcional y bioquímica de los componentes de la maquinaria de SUMO en Arabidopsis

En el sistema de la ubiquitina y SUMO, los enzimas y la cascada enzimática para la conjugación del polipéptido están conservados. Sin embargo, hay diferencias que hacen únicos estos sistemas, en el sistema de la ubiquitina existe tan sólo un modificador post-traduccional, decenas de enzimas E2 y cientos de E3 ligasas de la Ubiquitina. Son por tanto las E2 y E3 de la ubiquitina las que mayoritariamente confieren selectividad y especificidad de sustrato (Gareau and Lima, 2010). En cambio, en el sistema de SUMO, en plantas y vertebrados, existen diferentes isoformas de SUMO, solamente una enzima E2 conjugadora de SUMO y unas pocas E3 ligasas. El aumento de las isoformas de SUMO presentes en estos organismos sugiere características y funciones específicas para cada una de ellas.

En este contexto se han analizado las propiedades moleculares de los parálogos de SUMO en *A.thaliana*, *At*SUMO1, 2, 3 y 5 (capítulo I). Este estudio se ha centrado en el análisis de las interacciones no covalentes que media SUMO con la enzima activadora y conjugadora de SUMO y el motivo SIM que contienen algunas proteínas, características intrínsecas del sistema de sumoilación. Mientras que *Hs*SUMO1, 2 y 3 y *Smt3* presentan un alto grado de conservación de los residuos implicados en estas interacciones, los parálogos de SUMO de *Arabidopsis* presentan una superficie de interacción más divergente. Los resultados revelan que *At*SUMO1 y 2 son las isoformas que presentan una conservación más alta en los residuos implicados en las interacciones no covalentes con la E1 y la E2 de SUMO, lo que correlaciona con la mayor tasa de conjugación y capacidad de formación de cadenas de *At*SUMO1 y 2 respectivamente, mientras que *At*SUMO3 y 5 son las menos conservadas y presentan una menor tasa de conjugación. Esto sugiere que la mayor conjugación de *At*SUMO1 y 2 está determinada por el papel de la enzima activadora de SUMO en la discriminación entre los diferentes parálogos de SUMO.. Nos queda por elucidar si las

diferencias observadas en la superficie responsable de la interacción con los motivo SIM de los parálogos de SUMO tienen implicaciones a nivel de especificidad de sustrato u otras enzimas con motivo SIM que participan en la vía de SUMOilación, como *At*SIZ1, que podrían favorecer la conjugación de determinadas isoformas (Elrouby, 2014). Los resultados obtenidos concuerdan con el mayor papel de *At*SUMO1 y 2 *in vivo*. Se ha descrito el papel esencial de *At*SUMO1 y 2 para el correcto desarrollo de la planta (Saracco et al., 2007). Además, se han identificado sustratos de *At*SUMO1/2 y *At*SUMO3 en planta, mientras que la conjugación de *At*SUMO5 sólo ha sido detectada en plantas que sobrexpresan esta isoforma constitutivamente (Budhiraja et al., 2009; Kurepa, 2002). Asimismo, se han identificado proteasas específicas que reconocen *At*SUMO1 y 2 pero ninguna que reconozca eficientemente *At*SUMO3 y 5 (Chosed et al., 2006; Colby et al., 2006), lo que puede constituir un mecanismo para asegurar la maduración de las isoformas esenciales de SUMO in vivo. Por otro lado, solamente *At*SUMO1 y 2 son capaces de formar cadenas de SUMO lo que concuerda con la capacidad de estas dos isoformas de interaccionar con la enzima conjugadora de SUMO (Castaño Miquel et al., 2011; Colby et al., 2006). El hecho de que *At*SUMO3 no sea capaz de interaccionar con la enzima conjugadora de SUMO, sugiere que esta isoforma no pueda actuar como terminador de cadenas de SUMO. Todo esto apunta a un papel predominante de las isoformas *At*SUMO1 y 2 que está favorecido por el reconocimiento por parte de la enzima activadora y conjugadora de SUMO.

No solamente encontramos un aumento de diversidad en las isoformas de SUMO en *Arabidopsis*, curiosamente existen dos isoformas de la enzima activadora de SUMO, *At*E1a y *At*E1b, que difieren en la composición de la subunidad pequeña, *At*SAE1a y SAE1b. En *A.thaliana At*SAE1a está codificado por un único gen, mientras que para *At*SAE1b existen 2 copias génicas exactas, *At*SAE1b1 y *At*SAE1b2, que se encuentran duplicados en tándem en el genoma (capítulo II). La diversificación de la subunidad pequeña de la enzima activadora no es exclusiva de *A.thaliana* y se encuentra en todos los miembros de la especie *brassicaceae*, exceptuando *B.rapa*. En cambio, en levadura y mamíferos solamente existe una isoforma para la enzima activadora de SUMO. *At*E1a y *At*E1b presentan un 80% de identidad de secuencia, no obstante mediante ensayos *in vitro* hemos detectado una mayor actividad y estabilidad de la isoforma E1a. El estudio de mutantes *atsae1a* ha puesto de manifiesto que la enzima *At*E1a es necesaria para el mantenimiento de la homeostasis de SUMO en respuesta a estrés térmico y hídrico, condiciones en las que se induce una acumulación de conjugados de SUMO (Kurepa, 2002). Sin embargo, no hemos podido estudiar el papel de *At*SAE1b *in vivo,* porque se encuentra duplicado en tándem en el genoma de *Arabidopsis*. Con el fin de

esclarecer el papel de SAE1b en plantas, se podría estudiar la función *in vivo* de SAE1b en otros ecotipos de *Arabidopsis* que tan sólo contienen un gen que codifica para esta isoforma.

El estudio de la localización celular de la enzima activadora de SUMO, ha determinado que ambos heterodímeros presentan una localización nuclear, dependiente de una señal de localización nuclear clásica presente en el extremo C-terminal de SAE2, NLS1. Curiosamente, *At*SAE1a está presente en el citoplasma y núcleo de la célula. Nuestros resultados, muestran que *At*SAE1a podría existir como forma libre en el citoplasma de la célula, sugiriendo que en concentraciones limitantes de SAE2 se favorece la formación del heterodímero *At*E1a, más estable, activo e importante para el mantenimiento de la homeostasis de SUMO. Estos resultados apuntan a la existencia de mecanismos de regulación de la localización de los enzimas de la maquinaria de SUMO, en especial la enzima activadora. En humanos, se han identificado NLS para la subunidad grande, Uba2, y pequeña, Aos1, no obstante la importación del heterodímero está mediada por el reconocimiento de las importinas α/β de la señal bipartita de localización nuclear del extremo C-terminal de Uba2 (Moutty et al., 2011). La regulación de la localización de *Hs*E1 es también dependiente de la SUMOilación del extremo C-terminal de Uba2, que garantiza la retención de esta enzima en el núcleo de la célula (Truong et al., 2012a). Recientemente, en mamíferos y levaduras se ha puesto de manifiesto la importancia de diferentes modificaciones post-traduccionales en la regulación de los componentes de la maquinaria de SUMO, afectando a su localización, estabilidad y actividad. Apoyando esta línea de resultados, hemos identificado dos modificaciones post-traduccionales en el extremo C-terminal de SAE2 (capítulo V). Por un lado, se ha identificado durante el estadio de silicua el procesamiento de ocho aminoácidos del extremo C-terminal de SAE2, que constituyen una señal de localización nuclear no canónica, NLS2. El análisis de conservación del extremo C-terminal de SAE2 de las angiospermas y embriofitas muestra un alto grado de divergencia aunque hay ciertas zonas altamente conservadas, entre ellas se encuentran dos regiones básicas correspondientes a las NLS identificadas por nuestro grupo, NLS1 y NLS2, sugiriendo que este procesamiento post-traduccional podría estar conservado en las angiospermas y embriofitas. La localización de *At*E1 está regulada por dos señales independientes de localización nuclear presentes en el extremo C-terminal, siendo la NLS1 la que muestra una mayor contribución aunque no suficiente para la completa retención nuclear de la enzima. Postulamos que este procesamiento específico del extremo C-terminal de SAE2 puede ser un mecanismo de regulación de la conjugación de SUMO mediante el control de la

disponibilidad de la enzima activadora de SUMO en condiciones fisiológicas específicas. La identificación del mecanismo proteolítico por el cuál se procesa la NLS2 es crucial y nos ayudará a esclarecer cómo, cuándo y porqué sucede. No descartamos que este procesamiento tenga lugar en más etapas de desarrollo o en condiciones específicas de estrés, en las que sea necesaria la localización de SAE2 al citoplasma con el fin de favorecer la conjugación de sustratos específicos citoplasmáticos.

Por otro lado, mediante ensayos *in vitro* hemos identificado diferentes lisinas de SAE2, SAE1a, SUMO1 y SCE1 que son modificadas por SUMO (capítulo V). Varios de los sitios de SUMOilación analizados han sido identificadas *in vivo* por otros grupos (Miller and Vierstra, 2010). Se ha descrito que en respuesta a estrés térmico la SUMOilación de SAE2 y SCE1 decrece mientras que la forma SUMOilada de SIZ1 aumenta substancialmente, sugiriendo un bucle de autorregulación dentro de los miembros del sistema de SUMOilación (Miller and Vierstra, 2010). Además, modificaciones similares han sido identificadas en la sistema de SUMOilación de levadura y mamíferos, esta modificación puede dar lugar a una re-localización de la enzima, diferente especificidad de sustrato, interferir en interacciones con otras proteínas o promover la degradación de la enzima modificada vía proteosoma (Knipscheer et al., 2008; Truong et al., 2012a; 2012b). Todo esto sugiere que la SUMOilación de los componentes de la maquinaria de SUMO por sí mismos podría constituir un mecanismo conservado de regulación de la actividad, estabilidad y localización de estas enzimas *in vivo*. En especial, la actividad de la enzima E1 activadora de SUMO podría estar regulada por múltiples mecanismos moleculares, sugiriendo que constituye una etapa crítica in vivo.

Caracterización de nuevas implicaciones biológicas de SUMO en plantas

Una de las mayores limitaciones en el estudio de la SUMOilación *in vivo* es la obtención de plantas con la SUMOilación afectada, debido a la letalidad de los mutantes *atsumo1/atsumo2, atsce1 y atsae2* (Saracco et al., 2007). El efecto de SUMO *in vivo* se ha estudiado mayoritariamente mediante el estudio de las proteasas y mutantes de las E3 ligasas de SUMO (Hermkes et al., 2010; Murtas, 2003). Defectos en la vía de señalización por SUMO se han asociado a una mayor adaptación de la planta frente a algunos estreses, cómo la sequía, salinidad y mayor resistencia frente *PstDc3000* (Lee et al., 2006b; Miura et al., 2012; 2011), factores que causan grandes pérdidas económicas en los cultivos agrícolas. En este contexto, hemos generado una estrategia para inhibir la SUMOilación *in vivo*, basada en

la interacción entre E1^{UFD}-E2 que se encuentra conservada en la cascada enzimática de *las Ubiquitin like proteins* (Schulman and Wade Harper, 2009) (capítulo III). Además, se ha descrito en humanos y levadura el papel del dominio E1^{UFDCt} en la interacción no covalente con la enzima E2, requerido para la transesterificación de SUMO desde la enzima activadora a la conjugadora de SUMO (Lois and Lima, 2005). Mediante ensayos *in vitro* e *in vivo* hemos demostrado la capacidad de este dominio de interaccionar no covalentemente con la enzima conjugadora de SUMO en *A.thalina,* disminuyendo la conjugación de SUMO mediante la competición entre el dominio SAE2^{UFDCt} y la enzima E1 para interaccionar con la enzima conjugadora de SUMO. El dominio SAE2^{UFDCt} interacciona con el dominio LHB2 del E2, motivo divergente, sugiriendo que las interacciones E1-E2 son específicos para cada especie. La sobreexpresión del dominio SAE2^{UFDCt} en *A.thaliana* actúa como un dominante negativo, disminuyendo la capacidad de conjugación del sistema. Estas plantas muestran un fenotipo característico de plantas que presentan defectos en la conjugación de SUMO. En humanos, la modificación por SUMO se ha relacionado con una amplia variedad de enfermedades, incluyendo el cáncer, por ello se está estudiando como una potencial diana terapéutica. Se han identificado diferentes productos naturales y péptidos sintéticos que inhiben la conjugación de SUMO interaccionando con alguna de las enzimas de esta vía (E1, E2 o E3) (Fukuda et al., 2009a; 2009b; Hirohama et al., 2013; Kumar et al., 2013a; 2013b; Lu et al., 2010). La inhibición de la SUMOilación mediante la disrupción de dos enzimas (E1-E2) ofrece un aumento de especificidad y afinidad, representando un punto de partida para el desarrollo y optimización de péptidos inhibidores de esta interacción en otras especies, debido a la alta conservación del funcionamiento del sistema de SUMOilación entre especies.

La generación en el laboratorio de plantas con la SUMOilación afectada nos ha permitido estudiar el papel de SUMO en diferentes procesos biológicos, concretamente la respuesta de la planta frente la infección de patógenos (capítulo IV). La respuesta frente a patógenos está íntimamente relacionada con la señalización de las hormonas cómo el SA, JA y etileno (Glazebrook, 2005), y se ha descrito el papel de SUMO en el control de la homeostasis de SA (Lee et al., 2006b; van den Burg et al., 2010; Villajuana-Bonequi et al., 2014). Además, se ha descrito la acción de diferentes efectores de patógenos como proteasas de SUMO, afectando a la homeostasis de conjugados de SUMO en el huésped (Chosed et al., 2007; Hotson and Mudgett, 2004; Roden et al., 2004). Hasta el momento, se han desarrollado pocos estudios en relación con la SUMOilación y la respuesta a patógenos necrótrofos, que causan enfermedades de grandes pérdidas económicas en diferentes cultivos. Mediante el uso

de plantas con la SUMOilación afectada hemos estudiado el papel de SUMO frente a patógenos necrótrofos. Los resultados han mostrado que es necesario un correcto funcionamiento del sistema de SUMOilación para una activación de la respuesta de defensa de la planta frente a estos patógenos. Se ha descrito que el mutante *atsiz1-3* presenta resistencia frente al patógeno biótrofo bacterial *PstDc*3000, mientras que nuestros resultados muestran una mayor sensibilidad de este mutante frente la infección por los patógenos necrótrofos *Botrytis Cinerea* y *Plectospharaella cucumerina*. Está bien establecido que la resistencia del mutante *atsiz1-3* viene dada por un aumento de SA en la planta que activa de manera constitutiva la respuesta SAR (Lee et al., 2006a), indicando un diferente mecanismo de defensa de la planta según la clase de patógeno, la respuesta frente patógenos necrótrofos depende de una vía de señalización independiente al SA. Se ha descrito la importancia de las vías de señalización por JA y etileno frente patógenos necrótrofos, que activan la expresión de genes de defensa tales cómo *PDF1.2* y *Thi2.1* (John G Turner, 2002). Por tanto, sería interesante estudiar la vía de señalización por JA y Et en plantas con la SUMOilación afectada tanto en condiciones basales como la inducción de estás cascadas de señalización en respuesta a la infección. Además, no podemos descartar el efecto de ciertos efectores patógenos dañando la respuesta del huésped mediante la disminución del nivel de conjugados de SUMO en la célula. Que se podría dar tanto por efectores con actividad proteasa o señalización vía proteosoma de alguno de los componentes esenciales del sistema de SUMOilación. Con el fin de abordar esta posibilidad, actualmente en el laboratorio se está llevando a cabo el análisis de los niveles de proteína de conjugados de SUMO, E1 y E2 frente la infección por patógenos necrótrofos. Esperamos que futuros estudios en el laboratorio permitan establecer los principios básicos que se establecen entre los patógenos y el sistema de SUMOilación celular del huésped, que juega un papel esencial en la respuesta de la defensa de la planta.

Aparte de la implicación de SUMO en la respuesta frente a patógenos necrótrofos, otro estudio nos ha permitido establecer el papel de SUMO en la regulación de la ruptura de la dormancia (capítulo VI). La dormancia es una característica intrínseca de las plantas que controla la germinación de la planta en condiciones óptimas, de luz, humedad y temperatura. Hemos observado una mayor acumulación de conjugados de SUMO en semilla dormante, la cual decrece en condiciones no-dormantes. Con el fin de profundizar en la diferente dinámica de SUMOilación en ambas condiciones, se llevó a cabo una purificación de conjugados de SUMO en semillas secas (dormantes) y estratificadas (no dormantes), mediante plantas sobreexpresoras de HIS_6-SUMO, en condiciones desnaturalizantes. El procedimiento de

purificación de conjugados fue optimizado para incrementar la identificación de potenciales sustratos de SUMO y disminuir el número de falsos positivos. Además, como control se utilizaron semillas de plantas silvestres, Columbia. Las fracciones obtenidas, enriquecidas en conjugados, fueron separadas por HPLC y analizadas por espectroscopia de masas. Se siguió un criterio de selección de potenciales sustratos de SUMO, que resulto en la identificación de un total de 52 putativos sustratos, 27 específicos de semilla dormante y 22 de semilla no dormante. Curiosamente, tan sólo se identificaron 3 potenciales sustratos en ambas condiciones, estos resultados sugieren una SUMOilación especifica de sustratos en cada una de las condiciones analizadas. Por el contrario, estudios anteriores determinaron que en respuesta a estrés térmico, mayoritariamente incrementa la abundancia de conjugados ya existentes, más que la modificación de nuevos sustratos (Miller et al., 2013). Algunos de los sustratos identificados habían sido descritos anteriormente, sugiriendo que el método utilizado puede ser adaptado para purificar conjugados de SUMO o UBLs en diferentes condiciones. Los putativos sustratos identificados presentan tanto una localización nuclear como fuera de él, estos resultados apoyan la noción de que la SUMOilación tiene un papel como modificador post-traduccional fuera del núcleo.

CONCLUSIONES

1. **La etapa de activación de SUMO catalizada por la enzima heterodimérica SAE2/SAE1 tiene un papel regulador en la conjugación de SUMO.**

   1.1. Existe una conjugación preferencial de las isoformas esenciales *At*SUMO1/2 que está modulada por el papel de la enzima activadora en la discriminación de los parálogos de SUMO.

   1.2. En *Arabidopsis*, existen dos isoformas de la subunidad pequeña de la enzima activadora, SAE1a y SAE1b, que modulan de manera diferencial la tasa de conjugación de SUMO. El estudio de plantas mutantes que carecen de la subunidad SAE1a presentan deficiencias en la conjugación de SUMO, sugiriendo que la subunidad pequeña de la enzima activadora podría tener un papel regulador en el mantenimiento de la homeostasis de SUMO *in vivo*.

2. **La expresión del dominio UFDCt de la enzima activadora, SAE2^{UFDCt} inhibe la SUMOilación *in vivo* mediante un mecanismo de competencia por la unión de la enzima E2 conjugadora, bloqueando la transferencia de SUMO desde la enzima E1 a la E2.**

   2.1. Plantas que expresan el dominio SAE2^{UFDCt} mimetizan los defectos de desarrollo característicos de plantas que presentan una SUMOilación disminuida, validando esta estrategia para el estudio de la función de SUMO en plantas *in vivo*.

   2.2. Mediante el uso de estas plantas, se ha establecido que un sistema de SUMOilación funcional es necesario para la activación de las respuestas de defensa de las plantas frente patógenos necrótrofos.

3. **La enzima activadora de SUMO está regulada por modificaciones post-traduccionales.**

   3.1. Se ha identificado el procesamiento de ocho aminoácidos del extremo C-terminal de SAE2 que participan en la regulación de la localización subcelular de la enzima activadora. Mientras que la forma completa de SAE2 localiza exclusivamente en el núcleo celular, la forma procesada también está presente en el citoplasma. Este procesamiento, que se ha identificado de manera predominante durante el desarrollo

de semilla, podría constituir un mecanismo de regulación de la compartimentalizción de la SUMOilación in vivo.

3.2. Se ha identificado la SUMOilación de diferentes componentes de la maquinaria de SUMO *in vitro*, entre los que se encuentran SAE2, SAE1, SUMO y SCE1. Modificaciones similares en la maquinaria de SUMOilación de levadura y mamíferos han sido identificadas por otros investigadores, sugiriendo que estas modificaciones podrían constituir un mecanismo conservado de regulación de la actividad, la estabilidad y la localización de estas enzimas *in vivo*.

4. La SUMOilación tiene un papel importante durante las últimas etapas de formación de la semilla y en las primeras de su germinación.

4.1. Se ha observado que la acumulación de conjugados de SUMO en la semilla seca disminuye de manera dramática con el proceso de estratificación, sugiriendo que la SUMOilación juega un papel relevante en esta transición de desarrollo y que podría contribuir al mantenimiento de la semilla en estado de latencia.

4.2. El análisis proteómico de las fluctuaciones de la acumulación de conjugados de SUMO ha resultado en la identificación de un total de 52 putativos sustratos de SUMO en semilla seca y estratificada, aunque tan sólo 3 son comunes en ambas condiciones, sugiriendo una modificación específica de sustratos de SUMO durante la ruptura de la dormancia.

BIBLIOGRAFÍA

Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., and Becker, J. (1998). Structure determination of the small ubiquitin-related modifier SUMO-1. Journal of Molecular Biology *280*, 275–286.

Bernier-Villamor, V., Sampson, D.A., Matunis, M.J., and Lima, C.D. (2001). Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1. Cell *108*, 345–356.

Budhiraja, R., Hermkes, R., Muller, S., Schmidt, J., Colby, T., Panigrahi, K., Coupland, G., and Bachmair, A. (2009). Substrates Related to Chromatin and to RNA-Dependent Processes Are Modified by Arabidopsis SUMO Isoforms That Differ in a Conserved Residue with Influence on Desumoylation. Plant Physiology *149*, 1529–1540.

Bylebyl, G.R., Belichenko, I., and Johnson, E.S. (2003). The SUMO Isopeptidase Ulp2 Prevents Accumulation of SUMO Chains in Yeast. Journal of Biological Chemistry.

Castaño Miquel, L., Seguí, J., and Lois, L.M. (2011). Distinctive properties of ArabidopsisSUMO paralogues support the in vivopredominant role of AtSUMO1/2 isoforms. Biochem. J. *436*, 581–590.

Castro, P.H., Tavares, R.M., Bejarano, E.R., and Azevedo, H. (2012). SUMO, a heavyweight player in plant abiotic stress responses. Cell. Mol. Life Sci. *69*, 3269–3283.

Catalá, R., Ouyang, J., Abreu, I.A., Hu, Y., Seo, H., Zhang, X., and Chua, N.-H. (2007). The Arabidopsis E3 SUMO ligase SIZ1 regulates plant growth and drought responses. Plant Cell *19*, 2952–2966.

Chen, C.-C., Chen, Y.-Y., Tang, I.-C., Liang, H.-M., Lai, C.-C., Chiou, J.-M., and Yeh, K.-C. (2011). Arabidopsis SUMO E3 ligase SIZ1 is involved in excess copper tolerance. Plant Physiology *156*, 2225–2234.

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

Chosed, R., Tomchick, D.R., Brautigam, C.A., Mukherjee, S., Negi, V.S., Machius, M., and Orth, K. (2007). Structural analysis of Xanthomonas XopD provides insights into substrate specificity of ubiquitin-like protein proteases. Journal of Biological Chemistry *282*, 6773–6782.

Cohen-Peer, R., Schuster, S., Meiri, D., Breiman, A., and Avni, A. (2010). Sumoylation of Arabidopsis heat shock factor A2 (HsfA2) modifies its activity during acquired thermotholerance. Plant Mol Biol *74*, 33–45.

Colby, T., Matthaei, A., Boeckelmann, A., and Stuible, H.-P. (2006). SUMO-conjugating and SUMO-deconjugating enzymes from Arabidopsis. Plant Physiology *142*, 318–332.

Conti, L., Price, G., O'Donnell, E., Schwessinger, B., Dominy, P., and Sadanandom, A. (2008). Small Ubiquitin-Like Modifier Proteases OVERLY TOLERANT TO SALT1 and -2 Regulate Salt Stress Responses in Arabidopsis. The Plant Cell *20*, 2894–2908.

Depaux, A., Regnier-Ricard, F., Germani, A., and Varin-Blank, N. (2007). A crosstalk between hSiah2 and Pias E3-ligases modulates Pias-dependent activation. Oncogene *26*, 6665–6676.

Elrouby, N., and Coupland, G. (2010). Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify Arabidopsis proteins implicated in diverse biological processes. Proc. Natl. Acad. Sci. U.S.a. *107*, 17415–17420.

Elrouby, N. (2014). Extent and significance of non-covalent SUMO interactions in plant development. Plant Signal Behav *9*.

Fukuda, I., Ito, A., Hirai, G., Nishimura, S., Kawasaki, H., Saitoh, H., Kimura, K.-I., Sodeoka, M., and Yoshida, M. (2009a). Ginkgolic acid inhibits protein SUMOylation by blocking formation of the E1-SUMO intermediate. Chemistry & Biology *16*, 133–140.

Fukuda, I., Ito, A., Uramoto, M., Saitoh, H., Kawasaki, H., Osada, H., and Yoshida, M. (2009b).

Kerriamycin B inhibits protein SUMOylation. J. Antibiot. *62*, 221–224.

Garcia-Dominguez, M., March-Diaz, R., and Reyes, J.C. (2008). The PHD Domain of Plant PIAS Proteins Mediates Sumoylation of Bromodomain GTE Proteins. Journal of Biological ….

Gareau, J.R., and Lima, C.D. (2010). The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat Rev Mol Cell Biol *11*, 861–871.

Glazebrook, J. (2005). Contrasting mechanisms of defense against biotrophic and necrotrophic pathogens. Annu Rev Phytopathol *43*, 205–227.

Goldstein, G., Scheid, M., Hammerling, U., Schlesinger, D.H., Niall, H.D., and Boyse, E.A. (1975). Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. Proceedings of the ….

Hardeland, U., Steinacher, R., Jiricny, J., and Schär, P. (2002). Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. Embo J. *21*, 1456–1464.

Hay, R.T. (2005). SUMO: a history of modification. Molecular Cell *18*, 1–12.

Hermkes, R., Fu, Y.-F., Nürrenberg, K., Budhiraja, R., Schmelzer, E., Elrouby, N., Dohmen, R.J., Bachmair, A., and Coupland, G. (2010). Distinct roles for Arabidopsis SUMO protease ESD4 and its closest homolog ELS1. Planta *233*, 63–73.

Hershko, A., Ciechanover, A., and Varshavsky, A. (2000). The ubiquitin system. Nat. Med. *6*, 1073–1081.

Hirohama, M., Kumar, A., Fukuda, I., Matsuoka, S., Igarashi, Y., Saitoh, H., Takagi, M., Shin-ya, K., Honda, K., Kondoh, Y., et al. (2013). Spectomycin B1 as a novel SUMOylation inhibitor that directly binds to SUMO E2. ACS Chem. Biol. *8*, 2635–2642.

Hotson, A., and Mudgett, M.B. (2004). Cysteine proteases in phytopathogenic bacteria: identification of plant targets and activation of innate immunity. Current Opinion in Plant Biology *7*, 384–390.

Huang, L., Yang, S., Zhang, S., Liu, M., Lai, J., Qi, Y., Shi, S., Wang, J., Wang, Y., Xie, Q., et al. (2009). The Arabidopsis SUMO E3 ligase AtMMS21, a homologue of NSE2/MMS21, regulates cell proliferation in the root. The Plant Journal *60*, 666–678.

Ishida, T., Fujiwara, S., Miura, K., Stacey, N., Yoshimura, M., Schneider, K., Adachi, S., Minamisawa, K., Umeda, M., and Sugimoto, K. (2009). SUMO E3 Ligase HIGH PLOIDY2 Regulates Endocycle Onset and Meristem Maintenance in Arabidopsis. Plant Cell *21*, 2284–2297.

Ishida, T., Yoshimura, M., Miura, K., and Sugimoto, K. (2011). MMS21/HPY2 and SIZ1, Two Arabidopsis SUMO E3 Ligases, Have Distinct Functions in Development. PLoS ONE *7*, e46897–e46897.

Jin, J.B., Jin, Y.H., Lee, J., Miura, K., Yoo, C.Y., Kim, W.-Y., Van Oosten, M., Hyun, Y., Somers, D.E., Lee, I., et al. (2008). The SUMO E3 ligase, AtSIZ1, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through affects on FLC chromatin structure. Plant J. *53*, 530–540.

John G Turner, C.E.A.D. (2002). The Jasmonate Signal Pathway. The Plant Cell *14*, s153.

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

Johnson, E.S., and Gupta, A.A. (2000). An E3-like Factor that Promotes SUMO Conjugation to the Yeast Septins. Cell *106*, 735–744.

Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006). Modification of Proteins by Ubiquitin and Ubiquitin-Like Proteins. Annu. Rev. Cell Dev. Biol. *22*, 159–180.

Knipscheer, P., Flotho, A., Klug, H., Olsen, J.V., van Dijk, W.J., Fish, A., Johnson, E.S., Mann, M.,

Sixma, T.K., and Pichler, A. (2008). Ubc9 sumoylation regulates SUMO target discrimination. Molecular Cell *31*, 371–382.

Knipscheer, P., van Dijk, W.J., Olsen, J.V., Mann, M., and Sixma, T.K. (2007). Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. Embo J. *26*, 2797–2807.

Kotak, S., Larkindale, J., Lee, U., Koskull-Döring, von, P., Vierling, E., and Scharf, K.-D. (2007). Complexity of the heat stress response in plants. Current Opinion in Plant Biology *10*, 310–316.

Kumar, A., Ito, A., Hirohama, M., Yoshida, M., and Zhang, K.Y.J. (2013a). Bioorganic & Medicinal Chemistry Letters. Bioorganic & Medicinal Chemistry Letters *23*, 5145–5149.

Kumar, A., Ito, A., Hirohama, M., Yoshida, M., and Zhang, K.Y.J. (2013b). Identification of sumoylation activating enzyme 1 inhibitors by structure-based virtual screening. J Chem Inf Model *53*, 809–820.

Kurepa, J. (2002). The Small Ubiquitin-like Modifier (SUMO) Protein Modification System in Arabidopsis. ACCUMULATION OF SUMO1 AND -2 CONJUGATES IS INCREASED BY STRESS. Journal of Biological Chemistry *278*, 6862–6872.

Lee, C.-S., Chien, C.-T., Lin, C.-H., Chiu, Y.-Y., and Yang, Y.-S. (2006a). Protein changes between dormant and dormancy-broken seeds ofPrunus campanulata Maxim. Proteomics *6*, 4147–4154.

Lee, I., and Schindelin, H. (2008). Structural Insights into E1-Catalyzed Ubiquitin Activation and Transfer to Conjugating Enzymes. Cell.

Lee, J., Nam, J., Park, H.C., Na, G., Miura, K., Jin, J.B., Yoo, C.Y., Baek, D., Kim, D.H., Jeong, J.C., et al. (2006b). Salicylic acid-mediated innate immunity in Arabidopsis is regulated by SIZ1 SUMO E3 ligase. The Plant Journal *49*, 79–90.

Li, S.-J., and Hochstrasser, M. (2003). The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. The Journal of Cell Biology *160*, 1069–1081.

Lin, D., Tatham, M.H., Bin Yu, Kim, S., Hay, R.T., and Chen, Y. (2002). Identification of a substrate recognition site on Ubc9. Journal of Biological Chemistry *277*, 21740–21748.

Ling, Y., Zhang, C., Chen, T., Hao, H., Liu, P., Bressan, R.A., Hasegawa, P.M., Jin, J.B., and Lin, J. (2012). Mutation in SUMO E3 ligase, SIZ1, disrupts the mature female gametophyte in Arabidopsis. PLoS ONE *7*, e29470.

Liu, M., Shi, S., Zhang, S., Xu, P., Lai, J., Liu, Y., Yuan, D., Wang, Y., Du, J., and Yang, C. (2014). SUMO E3 ligase AtMMS21 is required for normal meiosis and gametophyte development in Arabidopsis. BMC Plant Biol. *14*, 153.

Lois, L.M. (2003). Small Ubiquitin-Like Modifier Modulates Abscisic Acid Signaling in Arabidopsis. The Plant Cell *15*, 1347–1359.

Lois, L.M. (2010). Diversity of the SUMOylation machinery in plants. Biochem. Soc. Trans *38*, 60.

Lois, L.M., and Lima, C.D. (2005). Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. Embo J. *24*, 439–451.

Lopez-Molina, L., Mongrand, S., Kinoshita, N., and Chua, N.-H. (2003). AFP is a novel negative regulator of ABA signaling that promotes ABI5 protein degradation. Genes Dev *17*, 410–418.

Lu, X., Olsen, S.K., Capili, A.D., Cisar, J.S., Lima, C.D., and Tan, D.S. (2010). Designed Semisynthetic Protein Inhibitors of Ub/Ubl E1 Activating Enzymes. J. Am. Chem. Soc. *132*, 1748–1749.

Mahajan, R., Delphin, C., Guan, T., Gerace, L., and Melchior, F. (1997). A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. Cell *88*, 97–107.

Matunis, M.J., Coutavas, E., and Blobel, G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. The Journal of Cell Biology *135*, 1457–1470.

Maupin-Furlow, J.A. (2014). Prokaryotic Ubiquitin-Like Protein Modification. Annu. Rev. Microbiol. *68*, 155–175.

Miller, M.J., Scalf, M., Rytz, T.C., Hubler, S.L., Smith, L.M., and Vierstra, R.D. (2013). Quantitative Proteomics Reveals Factors Regulating RNA Biology as Dynamic Targets of Stress-induced SUMOylation in Arabidopsis. Molecular & Cellular Proteomics *12*, 449–463.

Miller, M., and Vierstra, R. (2010). Mass spectrometric identification of SUMO substrates provides insights into heat stress-induced SUMOylation in plants. Plant Signal Behav *6*, 130–133.

Miura, K., Rus, A., Sharkhuu, A., Yokoi, S., Karthikeyan, A.S., Raghothama, K.G., Baek, D., Koo, Y.D., Jin, J.B., Bressan, R.A., et al. (2005). The Arabidopsis SUMO E3 ligase SIZ1 controls phosphate deficiency responses. Proc. Natl. Acad. Sci. U.S.a. *102*, 7760–7765.

Miura, K., and Hasegawa, P.M. (2010). Sumoylation and other ubiquitin-like post-translational modifications in plants. Trends in Cell Biology *20*, 223–232.

Miura, K., Jin, J.B., Lee, J., Yoo, C.Y., Stirm, V., Miura, T., Ashworth, E.N., Bressan, R.A., Yun, D.-J., and Hasegawa, P.M. (2007). SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in Arabidopsis. Plant Cell *19*, 1403–1414.

Miura, K., Lee, J., Jin, J.B., Yoo, C.Y., Miura, T., and Hasegawa, P.M. (2009). Sumoylation of ABI5 by the Arabidopsis SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. Proc. Natl. Acad. Sci. U.S.a. *106*, 5418–5423.

Miura, K., Okamoto, H., Okuma, E., Shiba, H., Kamada, H., Hasegawa, P.M., and Murata, Y. (2012). SIZ1deficiency causes reduced stomatal aperture and enhanced drought tolerance via controlling salicylic acid-induced accumulation of reactive oxygen species in Arabidopsis. The Plant Journal *73*, 91–104.

Miura, K., Sato, A., Ohta, M., and Furukawa, J. (2011). Increased tolerance to salt stress in the phosphate-accumulating Arabidopsis mutants siz1 and pho2. Planta *234*, 1191–1199.

Moutty, M.C., Sakin, V., and Melchior, F. (2011). Importin / mediates nuclear import of individual SUMO E1 subunits and of the holo-enzyme. Molecular Biology of the Cell *22*, 652–660.

Murtas, G. (2003). A Nuclear Protease Required for Flowering-Time Regulation in Arabidopsis Reduces the Abundance of SMALL UBIQUITIN-RELATED MODIFIER Conjugates. The Plant Cell *15*, 2308–2319.

Nigam, N., Singh, A., Sahi, C., Chandramouli, A., and Grover, A. (2008). SUMO-conjugating enzyme (Sce) and FK506-binding protein (FKBP) encoding rice (Oryza sativa L.) genes: genome-wide analysis, expression studies and evidence for their involvement in abiotic stress response. Mol. Genet. Genomics *279*, 371–383.

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

Olsen, S.K., Capili, A.D., Lu, X., Tan, D.S., and Lima, C.D. (2010). Active site remodelling accompanies thioester bond formation in the SUMO E1. Nature *463*, 906–912.

Pichler, A., Gast, A., Seeler, J.S., Dejean, A., and Melchior, F. (2001). The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity. Cell *108*, 109–120.

Pichler, A., Knipscheer, P., Oberhofer, E., van Dijk, W.J., Körner, R., Olsen, J.V., Jentsch, S., Melchior, F., and Sixma, T.K. (2005). SUMO modification of the ubiquitin-conjugating enzyme E2-25K. Nat Struct Mol Biol *12*, 264–269.

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

Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001). SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. Journal of Biological Chemistry *276*, 12654–12659.

Rytinki, M.M., Kaikkonen, S., Pehkonen, P., Jääskeläinen, T., and Palvimo, J.J. (2009). PIAS proteins: pleiotropic interactors associated with SUMO. Cell. Mol. Life Sci. *66*, 3029–3041.

Saracco, S.A., Miller, M.J., Kurepa, J., and Vierstra, R.D. (2007). Genetic Analysis of SUMOylation in Arabidopsis: Conjugation of SUMO1 and SUMO2 to Nuclear Proteins Is Essential. Plant Physiology *145*, 119–134.

SCHLESINGER, D.H., and GOLDSTEIN, G. (1975). Molecular conservation of 74 amino acid sequence of ubiquitin between cattle and man. Nature *255*, 423–424.

Schulman, B.A., and Wade Harper, J. (2009). Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. Nat Rev Mol Cell Biol *10*, 319–331.

Sharrocks, A.D. (2006). PIAS proteins and transcriptional regulation--more than just SUMO E3 ligases? Genes Dev *20*, 754–758.

Song, J., Durrin, L.K., Wilkinson, T.A., Krontiris, T.G., and Chen, Y. (2004). Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. Proc. Natl. Acad. Sci. U.S.a. *101*, 14373–14378.

Sydorskyy, Y., Srikumar, T., Jeram, S.M., Wheaton, S., Vizeacoumar, F.J., Makhnevych, T., Chong, Y.T., Gingras, A.C., and Raught, B. (2010). A Novel Mechanism for SUMO System Control: Regulated Ulp1 Nucleolar Sequestration. Molecular and Cellular Biology *30*, 4452–4462.

Tatham, M.H., Geoffroy, M.C., Shen, L., and Plechanovova, A. (2008). RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation : Abstract : Nature Cell Biology. Nature Cell ….

Truong, K., Lee, T.D., and Chen, Y. (2012a). Small Ubiquitin-like Modifier (SUMO) Modification of E1 Cys Domain Inhibits E1 Cys Domain Enzymatic Activity. Journal of Biological Chemistry *287*, 15154–15163.

Truong, K., Lee, T.D., Li, B., and Chen, Y. (2012b). Sumoylation of SAE2 C Terminus Regulates SAE Nuclear Localization. Journal of Biological Chemistry *287*, 42611–42619.

Van den Burg, H.A., Kini, R.K., Schuurink, R.C., and Takken, F.L.W. (2010). Arabidopsis Small Ubiquitin-Like Modifier Paralogs Have Distinct Functions in Development and Defense. The Plant Cell *22*, 1998–2016.

Vertegaal, A. (2011). Uncovering ubiquitin and ubiquitin-like signaling networks. Chemical Reviews.

Vijay-Kumar, S., Bugg, C.E., and Cook, W.J. (1986). Structure of ubiquitin refined at 1.8Åresolution. Journal of Molecular Biology *194*, 531–544.

Villajuana-Bonequi, M., Elrouby, N., Nordström, K., Griebel, T., Bachmair, A., and Coupland, G. (2014). Elevated salicylic acid levels conferred by increased expression of ISOCHORISMATE SYNTHASE 1 contribute to hyperaccumulation of SUMO1 conjugates in the Arabidopsis mutant early in short days 4. The Plant Journal n/a–n/a.

Walden, H., Podgorski, M.S., and Schulman, B.A. (2003). Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8. Nature *422*, 330–334.

Wang, J., Hu, W., Cai, S., Lee, B., Song, J., and Chen, Y. (2007). The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications. Molecular Cell *27*, 228–237.

Wang, J., Taherbhoy, A.M., Hunt, H.W., Seyedin, S.N., Miller, D.W., Miller, D.J., Huang, D.T., and

Schulman, B.A. (2010). Crystal Structure of UBA2ufd-Ubc9: Insights into E1-E2 Interactions in Sumo Pathways. PLoS ONE *5*, e15805.

Watts, F.Z. (2013). Starting and stopping SUMOylation. Chromosoma *122*, 451–463.

Werner, A., Flotho, A., and Melchior, F. (2012). The RanBP2/RanGAP1∗SUMO1/Ubc9 Complex Is a Multisubunit SUMO E3 Ligase. Molecular Cell *46*, 287–298.

Wilkinson, K.A., and Henley, J.M. (2010). Mechanisms, regulation and consequences of protein SUMOylation. Biochem. J. *428*, 133–145.

Yoo, C.Y., Miura, K., Jin, J.B., Lee, J., Park, H.C., Salt, D.E., Yun, D.J., Bressan, R.A., and Hasegawa, P.M. (2006). SIZ1 Small Ubiquitin-Like Modifier E3 Ligase Facilitates Basal Thermotolerance in Arabidopsis Independent of Salicylic Acid. Plant Physiology *142*, 1548–1558.

AGRADECIMIENTOS

Per fi arriba el moment d'escriure els agraïments, sembla mentida, però al final aquí està la meva tesi. I no ho hauria pogut fer sense molta gent que m'ha ajudat i m'ha acompanyat, agafant-me de la mà, durant aquest camí llarg i sorprenent. En primera instancia, vull agrair a la meva supervisora de tesi la Dra. Maria Lois, o "Mari", la oportunitat que ja fa quasi 6 anys enrere, recent sortida de la carrera de bioquímica, em va oferir. Gràcies per haver confiat en mi, ha sigut una bona experiència amb constant aprenentatge.

Aprofito per donar les gràcies a la Dra. Maria Coca, per la seva paciència amb mi i la seva ajuda en el projecte que tenim en col·laboració, sempre disposada a donar un cop de mà. Gràcies al Dr. David Reverter que m'ha obert una mica les portes al món de la cristal·lografia. I per suposat, vull donar moltíssimes gràcies al Dr. Alfred Vertegaal, que em va acollir al seu laboratori, com si fos una més i es va preocupar per mi en tot moment, facilitant-me la meva estada a Leiden.

Durant aquests anys ha passat molta gent pel SUMO lab... Molts d'ells grans amics i amigues, quantes hores, moments, plors i riures, no sé ni però on començar. Però vaja, començaré donant les gràcies al Josep, el primer doctorant de la Mari, que em va ajudar molt durant el meu inici i gràcies a la seva "herència científica" puc avui presentar la meva tesi, sense ell no hagués pogut fer-ho tan fàcilment. Al Rubén, al Félix, l'Arnau, la Fagua i la Maria per tots els bons moments que hem compartit junts al CID i posteriorment al CRAG. I a les nenes, la Natalia, la Inesita i la Silvy, perquè hem fet sempre pinya i hem format molt bon equip, i pk sou fantàstiques cada una de vosaltres. Silvy, la millor post-doc que podríem haver tingut, sempre disposada a ajudar, sempre m'he preguntat d'on treus tota aquesta energia positiva, no se't acaben mai les piles!!! Natalia, hem compartit milions de coffee-breaks increïbles, la millor companyia en els moments de desesperació. Inês tu realidad supera la ficción, eres genial, como hecho de menos tus consejos! Suerte en tu tesi! Em va costar acostumar-me a la vostra absència... però per sort, seguim en contacte, ja sigui per e-mail, skype o wassap des de Barcelona amb connexió directe a Milán, Cambridge o Lisboa. Per sort, no em vaig quedar mai sola al laboratori i he pogut compartir els meus últims moments amb l'Arny i l'Abraham. Arny, mucha paciència y suerte en tu post-doc!! Y cuida de Abraham!!! Abraham, espero que tot et vagi molt bé i estic segura que faràs una súper tesi! Ànims

hi ha moments durs, però al final sempre son recompensats i val la pena! Qualsevol cosa, ja saps on estic!

Aquesta tesi no hagués sigut al mateix sense tota l'altre gent del CRAG; en especial el grup del Riechman amb qui hem compartit laboratori. Mariana mucha suerte con la tesi! Thilia, suerte teníamos de que pusieras orden en el laboratorio!!!! Jin, quizás te llamo para que me ofrezcas una plaza de post-doc en tu grupo ;) Tomyyyy, ha sido una sorpresa conocerte, y me has ayudado mucho. Aunque, no voy a negar que eres un pesado jajajajaja
A la Catalina i el Nico, per les nostres confessions, i per aquells dinars de diumenge a la Vall d'hebron amb la gent del grup del Manuel. A la Águila, la sevillana más dulce del centro! Heavy, gràcies pels teus bons consells i la teva bona fe. Lucio, Briardo, Miriam, Pablo, siempre ha sido un placer compartir con vosotros este tramo de mi vida. Mucha suerte, i nos vemos por BCN! A mi primo, el Manzano no me olvido de ti! Gracias por estar cerca, y ayudar cuando es necesario! Extiendo las gracias a compañeros que ya no están en el CRAG Tommaso, Belmiro, Nahuel, Isa, Jorge, Jud, Maria, Helene, Adrián ... y muchos más, gracias por los buenos momentos!
I no pot faltar, la meva primera i única companya de pis, la Guiomar. Ha estat un plaer compartir amb tu tan la vida al laboratori com la casera. Quants records... molta sort en el tram final! I la Nogmeee, com tu vas dir has sigut un dels meus nous descobriments al CRAG! Un dia d'aquets hem de fer les maletes i acaba el nostre viatge. Ens ho devem!

I com no, tota la gent que tot i no formar part del CRAG, formen part de la meva vida i m'han recolzat en tot moment. Les amigues de la universitat l'Agnes, l'Anna, la Laura i la Ivet, amb qui vaig compartir moltes hores de classe, laboratori, biblioteca i bar. Gràcies, per estar al meu costat tot aquest temps. I els amics de tota la vida: l'Anna, l'Esther, la Mireia, l'Alex, el Gerard, la Mònica i l'Alba. Sembla increïble que després de tants anys i lo diferent que som siguem sent amics, però es que tenim molts records compartits junts que ens uneixen cada dia més. I sí, Dr. Gonzalez i Dra. Oteo ara hi haurà una doctora més al grup!!!

I vull fer menció especial a la meva família, a la meva mare perquè sempre ha confiat en mi i m'ha incitat a ser arriscada i ha viure la vida al màxim, com fa ella. Al meu pare, perquè m'ha ensenyat que es pot aconseguir el que un vulgui però t'hi has d'esforçar al màxim, i fer les coses bé i amb delicadesa. Al meu germà, l'Oriol, perquè es un crack i la meva vida no seria igual sense ell; Em deus unes birres!!! A la meva estimada germana petita, Ariadna, que em fa veure la vida amb uns altres ulls, quan estic amb ella m'oblido de les meves preocupacions. Al Lluís perquè m'ha ensenyat a ser valenta, i a l'Àngels pels seus bons consells. I en particular, als meus avis, perquè em sento afortunada de poder compartir amb ells molts moments de la meva vida.

I finalment, al Albert, que m'ha acompanyat al final d'aquest tram. Gràcies, sense tu hagués estat molt més difícil, m'has ajudat a relativitzat les meves preocupacions i ha donar al màxim de mi mateixa. I el més important de tot, sempre m'has fet costat en cada una de les decisions que he anat prenent durant aquest camí. Sé que estàs fart de la paraula SUMO, entre moltes altres, però ja s'acaba i ara comença una nova etapa. I en tinc moltes ganes, no tinc clar que ens espera, però m'agrada tenir al davant un camp d'alternatives varies i tenir-te a tu al meu costat per fer junts el següent pas.

Gràcies a tots!