Spring 2015

Developmentally regulated SUMOylation in the ciliate Tetrahymena thermophila

Amjad M. Nasir
Purdue University

Graduate School Form 30
Updated 1/15/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Amjad M. Nasir

Entitled

DEVELOPMENTALLY REGULATED SUMOYLATION IN THE CILIATE TETRAHYMENA THERMOPHILA

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

James D. Forney (Chair)

Ourania M. Andrisani

Scott D. Briggs

Mark C. Hall

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): _____

Approved by: James D. Forney

Head of the Departmental Graduate Program Date

DEVELOPMENTALLY REGULATED SUMOYLATION IN THE CILIATE *TETRAHYMENA THERMOPHILA*

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Amjad M. Nasir

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

May 2015

Purdue University

West Lafayette, Indiana

وَٱللَّهُ خَلَقَ كُلَّ دَآبَّةٍ مِّن مَّآءٍ ۖ فَمِنْهُم مَّن يَمْشِى عَلَىٰ بَطْنِهِۦ وَمِنْهُم مَّن يَمْشِى عَلَىٰ رِجْلَيْنِ وَمِنْهُم مَّن يَمْشِى عَلَىٰٓ أَرْبَعٍ ۚ يَخْلُقُ ٱللَّهُ مَا يَشَآءُ ۚ إِنَّ ٱللَّهَ عَلَىٰ كُلِّ شَىْءٍ قَدِيرٌ ﴿٤٥﴾

TABLE OF CONTENTS

Page

LIST OF FIGURES iv

LIST OF TABLES v

ABSTRACT vii

CHAPTER 1 – INTRODUCTION 1

CHAPTER 2 – SUMOYLATION IS DEVELOPMENTALLY REGULATED AND REQUIRED FOR CELL PAIRING DURING CONJUGATION IN *TETRAHYMENA* 26

CHAPTER 3 – PROTEOMIC STRATEGY TO IDENTIFY SMT3P SUBSTRATES IN *TETRAHYMENA* 64

FUTURE DIRECTIONS 92

REFERENCES 100

APPENDIX 116

VITA 128

LIST OF FIGURES

Figure Page

Figure 1. SUMOylation occurs through a 3-step enzymatic pathway 8

Figure 2. The Sexual Life Cycle of *Tetrahymena* 23

Figure 3. Smt3p localizes to the somatic MAC during vegetative growth 39

Figure 4. Differential modification of proteins by Smt3p during conjugation versus vegetative growth 42

Figure 5. Localization of Uba2p and Smt3p in conjugation 45

Figure 6. Generation of *SMT3* and *UBA2* knockout heterokaryons 48

Figure 7. *Tetrahymena* expressing the *GFP-UBA2* or *FLAG-His6-SMT3* transgene regulated by the *MTT1* promoter behave as conditional mutants 53

Figure 8. Depletion of Smt3p and Uba2p results in inability to form mating pairs 56

Figure 9. Smt3p and Uba2p depleted cell lines are sensitive to DNA damaging agents 58

Figure 10. The psYFP2_His6-3FLAG-SUMO vector and its expression product 74

Figure 11. Overview of the double affinity purification and identification of His6-3FLAG-Smt3p tagged proteins 76

Figure 12. Comparison of SUMO substrates identified using 2-step purification in various model systems 79

Figure 13. Overlap between Smt3p substrates purified under denaturing and native conditions 80

Figure 14. Graphical representation of both the cellular location and biological function of Smt3p candidate substrates 81

LIST OF TABLES

Table Page

Table 1. Nomenclature of SUMO Pathway Enzymes in Model Systems 3

Table 2. *SMT3* and *UBA2* are essential genes 48

Table 3. List of selected Smt3p substrates identified in vegetative *Tetrahymena* 75

Table 4. Consensus Smt3p target sites 81

Table 5. Predicted frequency of Smt3p consensus sites of proteins identified by different studies 82

Table 6. List of oligos used in this study 115

Table 7. List of identified Smt3p protein substrates from vegetative *Tetrahymena* 117

ABSTRACT

Nasir, Amjad, Ph.D., Purdue University, May 2015. Developmentally Regulated SUMOylation in the ciliate *Tetrahymena thermophila.* Major Professor: James D. Forney.

The Small Ubiquitin-like MOdifier (SUMO) protein regulates numerous nuclear events such as transcription, mitosis and meiosis and DNA repair. These processes are critical to the programmed nuclear events of conjugation in ciliates and provide the potential to investigate developmentally regulated SUMOylation. We predicted a developmental increase in SUMOylation during late conjugation based on the extensive genome remodeling in the developing macronucleus (anlagen) of *Tetrahymena thermophila*. Immunoblotting of cell lysates from vegetative and mating cells using anti-SUMO antibodies revealed distinct developmental differences and an increased signal correlated with formation of the anlagen. Immunofluorescence of mating *Tetrahymena* cells with the same antibody revealed an increase in staining of the parental macronucleus until the signal shifts to the anlagen at 7 hours post-mixing. This along with the finding that GFP-Uba2 fusion proteins localize to the anlagen is consistent with a major nuclear role for

SUMOylation during conjugation. Germ-line knockout mutants of *SUMO* (*SMT3*) and *UBA2* are vegetative lethal and conditional mutants dependent on a cadmium-inducible metallothionein promoter exhibit reduced cell growth and increased sensitivity to DNA damaging agents upon cadmium withdrawal. Interestingly, mating of Uba2p conditional lines leads to a cadmium-dependent delay after meiosis but prior to macronuclear development predicting a SUMOylation-dependent event. Additionally, in an effort to provide further insight into the various processes affected by SUMO modification, we utilized a proteomics-based approach to generate an unbiased spectrum of SUMO protein substrates from vegetative *Tetrahymena*. We used a two-step affinity purification scheme to isolate SUMO substrates from a tagged *Tetrahymena* strain. Proteins from the purification were then identified by subsequent LC-MS.MS analysis using a QuadTOF mass spectrometer. We identified 110 candidate proteins that were identified by 2 or more peptides. The nature of the protein substrates that were identified is consistent with roles of SUMOylation in modulating diverse cellular processes including transcriptional regulation, protein folding and translation, metabolism and crosstalk with other post-translational modifications. These results support multiple roles of SUMOylation in regulating various cellular processes in eukaryotic cells. Taken together, our findings provide the foundation for additional studies of SUMOylation during conjugation in *Tetrahymena*.

CHAPTER 1 – INTRODUCTION

SUMO is a Member of the Ubiquitin-like Protein Family

Protein function is often regulated though additional or removal of molecular adducts collectively termed post-translational modification (PTM). Because of the dynamic nature of such modifications, PTMs are capable of rapid modulating protein function by affecting stability, localization as well as interaction with other proteins thereby altering the cellular proteome. Ubiquitin-like (UBL) or ubiquitin-related modifiers are post-translationally attached to substrate proteins via an enzymatic pathway reminiscent to that used in ubiquitin (Ub). SUMOylation is one such UBL modification which involves the reversible and covalent conjugation of members of the SUMO (Small Ubiquitin-like MOdifier) family of proteins to various target proteins. Even though SUMO belongs to the Ub-like family of modifiers, its functions are very diverse and different from Ub itself.

SUMO Isoforms and Structure

SUMOylation affects several critical processes in eukaryotic cells and is required for cell viability in most eukaryotic systems (reviewed in (1)). Simple eukaryotes such as yeast, fruit fly and nematodes express a single SUMO gene.

In yeast, the SUMO homolog is named Smt3 (Suppressor of Mif Two 3) and the same nomenclature will be used throughout this thesis when referring to the gene (*SMT3*) or its product (Smt3p) in *Tetrahymena* (Table 1). There are four isoforms of the SUMO protein in mammalian cells, with SUMO-2 and SUMO-3 (also abbreviated as SUMO-2/3) sharing 97% sequence identity with one another and only ~50% identity with SUMO-1 (reviewed in (2)). A fourth SUMO protein (SUMO-4) has been described in humans, and is expressed primarily in kidney cells (3). SUMO-1 shares the highest sequence similarity with SUMO in yeast and most lower eukaryotes and is the primary isoform that is conjugated on to protein substrates under normal conditions. SUMO2/3-adducts appear when cells undergo physiological stress although there are examples of proteins such as topoisomerase II, which disentangles DNA strands during replication, and CENP-E, which is a component of centromeres in mammals, that are modified by SUMO-2/3 under normal conditions too (4, 5). While there have been reports on paralog-specific functions of SUMO-2/3, these proteins display functional redundancy and have considerable overlap in substrate specificities with SUMO-1 (6, 7).

The structures of human (8) and yeast (9) SUMO variants have been resolved recently and tertiary structures of both Ub and SUMO are very similar and nearly superimposable even though SUMO proteins share very little (~18%) sequence similarity with Ub (Figure 1) (10). A distinguishing characteristic of SUMO is that it contains an elongated and flexible N-terminal neck region that has been linked to polySUMO chain formation (11, 12). Remarkably, yeast mutants that lack the

extended neck region including the lysine residues required for auto-SUMOylation are viable without deleterious phenotypes which would suggest that, unlike Ub, the formation of polymeric SUMO chains is not essential in *S. cerevisiae* (13). Another similarity between SUMO and Ub is that the mature forms of both proteins end with two glycine residues (diglycine motif) necessary for conjugation of these proteins to lysine residues on target proteins (reviewed in (14, 15)).

Table 1. Nomenclature of SUMO Pathway Enzymes in Model Systems.

Enzyme	**Humans/Mammals**	***S. cerevisiae***	***T. thermophila***
SUMO paralogs	SUMO-1 SUMO-2 SUMO-3 SUMO-4	Smt3p	Smt3p
E1 Activating Enzyme	SAE1/SAE2	Aos1p/Uba2p	Aos1p/Uba2p
E2 Conjugating Enzyme	UBC9	Ubc9p	Ubc9p
E3 Ligases	PIAS1	Siz1p	Unidentified
	PIAS3 PIASx PIASy	Siz2p or Nfi1p	Unidentified
	Mms21	Mms2p or Zip3p	
	Pc2	Unidentified	
	RanBAP	Unidentified	
SUMO Proteases	SENP1 SENP2 SENP3	Ulp1p	Unidentified
	SENP5	Ulp2p or Smt4p	Unidentified
	SENP6 SENP7		

SUMOylation in Human Disease

In the past few years, there have been increasing reports linking SUMOylation to human disease pathogenesis. Most notably, dysregulation of the SUMO pathway has been implicated in tumorigenesis and cancer onset (reviewed in (16, 17)). This is not surprising as SUMOylation regulates important tumor suppressor proteins such as p53, MDM2 and pRB (16). The link between SUMOylation and cancer is further solidified by several studies where increased levels of SUMO E1 Activating enzyme (18, 19), E2 Conjugating enzyme (20), SUMO E3 ligases (21) and SUMO proteases (22, 23) have been shown to contribute to human cancers. Myc is an oncogenic transcription factor frequently dysregulated in human cancer. In human breast cancer, SAE2 (Uba2p in *S. cerevisiae* and *T. thermophila*) is required for Myc-dependent tumorigenesis and patients with high levels of SAE2 suffer from increased metastasis (19). Inhibition of SUMOylation in this case may provide therapeutic benefits for patients with Myc-driven cancers.

Many proteins that play critical roles in neurodegenerative diseases have also been identified as targets of SUMO. This list includes, but is not limited to, huntingin (Huntington's disease), ataxin-1 (spinocerebellar ataxia type 1), tau protein (Parkinson's and Alzheimer's disease) and SOD1 (amyotrophic lateral sclerosis) (reviewed in (24)). Diabetes, which is a very common human disease, has also been linked to SUMOylation (25, 26). In mouse, the transcription factor c-Maf is a target of SUMO. SUMOylation of c-Maf reduces its transactivational capacity by sequestering it in promyelocytic leukemia nuclear bodies (PML-NBs). This prevents

c-Maf from binding to interleukin IL-4 which otherwise confers resistance to diabetic state (26). A similar study in mouse showed that overexpression of SUMO-2 negatively regulates the transcriptional factor NF-κB. The reduced activity of NF-κB confers protection against diabetes in transgenic mice overexpressing SUMO-2 as NF-κB is unable to activate genes involved in the development of Type-1 diabetes (reviewed in (27)). Another major human disease linked to SUMOylation is cardiomyopathy which most commonly causes heart failure. Overexpression of the SUMO protease, SENP5 has been shown to result in elevated apoptosis leading to cardiac dysfunction (28). The increase in cell apoptosis is also seen with overexpression of the SUMO conjugating enzyme, Ubc9p where such apoptotic events precede other detectable pathological changes suggesting its underlying role in cardiomyopathy (29). Understanding the role of SUMOylation in human disease is, therefore, of tremendous utility as it provides insights into functions of protein substrates and mechanistic pathways that, when perturbed, cause diseased states.

The SUMO Conjugation Pathway Involves 3 Enzymes

SUMO-Activating Enzyme

Modification of proteins by SUMO occurs through a biochemical pathway that is similar to attachment of Ub to target proteins (30). The first step in the pathway utilizes an E1-activating enzyme which is a heterodimer comprised of two subunits – SAE1 and SAE2 known as Aos1 and Uba2 in yeast respectively (31, 32) (Table 1 and Figure 1). It is interesting to note that Aos1 and Uba2 resemble the amino- and

carboxyl-termini of Ub E1 enzymes (6, 32–34). Activation of SUMO occurs through an ATP-dependent step in which the carboxyl group of the C-terminal glycine of SUMO forms a thioester bond with a cysteine residue in the active site of the Uba2 component of the E1 enzyme. In most eukaryotic organisms, there is a single E1 enzyme required for the activation of SUMO proteins (6). This is in stark contrast to the Ub system, where several E1 enzymes (8 in humans for example) are known to initiate Ub activation (35, 36). There is no overlap of function between SUMO E1 enzymes (as well as E2 and E3 enzymes) with the Ub system and vice-versa. In mammals and vertebrates, the single SUMO E1 enzyme will activate all isozymes of SUMO (34, 37). Budding yeast strains that lack Aos1 or Uba2 exhibit a lethal phenotype which indicates that the SUMO E1 is an essential enzyme in this organism (32, 33). In fission yeast, *S. pombe,* Aos1 deletion mutants are able to conjugate SUMO on to protein substrates at very low levels (38). This would suggest that Uba2 is the essential component of the E1 enzyme as it contains the catalytic active site. A less likely explanation would be that Aos1 paralogs from other UBL modifications may substitute for Aos1.

Uba2 is expressed in all stages of *Drosophila* life cycle with an increase during embryogenesis that would indicate a requirement for SUMOylation during this stage of the life cycle (39). In the ciliate *Tetrahymena* thermophila, *UBA2* transcripts are expressed throughout the vegetative life cycle but are most abundant during the sexual life cycle with a peak during the later stages corresponding to the generation, development and maturation of new somatic nuclei (40). This expression pattern is

consistent with the observation that E1 enzymes are found predominantly in the cell nuclei of other eukaryotic model systems (6, 33, 39).

SUMO-Conjugating Enzyme

In the second step of the SUMO conjugation pathway, activated Smt3p is transferred from the E1-activating enzyme to a cysteine residue in the active site of the E2-conjugating enzyme. Unlike the Ub system which displays a greater diversity of E2 enzymes, only a single E2 SUMO-conjugating enzyme, known as Ubc9, has been identified as of yet (7) (Table 1 and Figure 1). A second distinguishing feature of Ubc9 from other Ub E2 enzymes is that it is able to directly recognize substrate proteins and catalyze the formation of isopeptide bond between SUMO and substrate proteins in the absence of E3 ligases (reviewed in (2)).

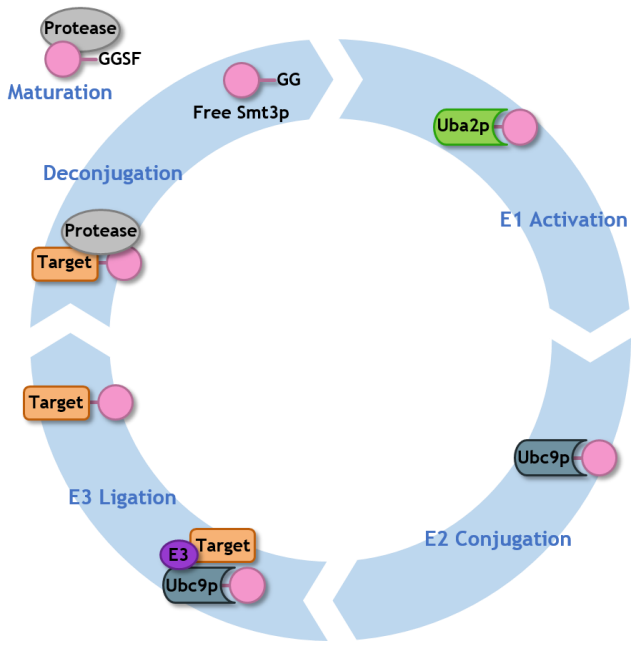

Figure 1. SUMOylation occurs through a 3-step enzymatic pathway. Conjugation of Smt3p to substrate proteins occurs through a 3-step cascade that shares numerous similarities to the ubiquitin pathway. Precursor Smt3p undergoes cleavage by SUMO proteases to produce mature Smt3p revealing a C-terminal diglycine motif. Mature Smt3p is then activated by the E1 Activating enzyme, which is a heterodimer of Uba2p and Aos1p, in an ATP-dependent manner (only Uba2p is shown here). The E2 conjugating enzyme, Ubc9p, assisted by several adapter proteins known as SUMO E3 ligases, will then conjugate Smt3p onto a Lys residue of the substrate protein. SUMO proteases are also responsible for removing Smt3p moiety from substrates generating free Smt3p in the process.

SUMO-Ligating Enzymes

In the final step of the SUMO pathway, the SUMO-E2 thioester intermediate serves as a donor in the conjugation of SUMO moiety onto a substrate protein. The terminal glycine in SUMO forms an isopeptide bond with the ϵ amino group of a lysine residue in the target protein. Three groups of proteins have been identified in recent years that accelerate the conjugation of SUMO onto substrate proteins. These proteins share very few recognizable features and are collectively known as SUMO

E3 ligases. Surprisingly, these SUMO E3 ligases bear no semblance to typical Ub E3 ligases. The most famous SUMO E3 enzymes are the Siz/PIAS family of proteins that contain the SP-RING that is analogous to RING domains of Ub E3 ligases. Siz/PIAS proteins bind to the E2-SUMO thioester intermediate and protein substrates which brings them in close proximity and promotes Smt3p transfer (reviewed in (2, 41)).

A second and distinct class of E3 enzymes is represented by RanBP2 that sits at the nuclear pore complex and aids in the SUMO-modification of its target RanGAP1. RanBP2 does not contain either RING or HECT domains that are found in Ub E3 ligases and its E3 ligase domain favors preferential modification of RanGAP with SUMO-1 over modification by SUMO-2 in mammalian systems (41). RanBP2 does not recruit RanGAP1 to Ubc9 but rather alters the structure of the SUMO-1-Ubc9 thioester intermediate such that it has increased capacity to transfer SUMO to its RanGAP1 substrate.

The last group of SUMO ligases comprise the human polycomb group protein 2 (Pc2) which is a member of a large multi-protein complex that regulates transcriptional repression by altering chromatin structure. The effect of Pc2 in the increased SUMOylation of its substrate CtBP2 is modest and direct interactions between Pc2 and the SUMO-Ubc9 thioester intermediate have not been noted (reviewed in (41)).

Substrate Selection for SUMOylation

Attachment of SUMO to target proteins occurs on lysine residues that are typically part of the SUMO-modification consensus motif, ψKXE (where ψ represents aliphatic amino acids such as leucine, isoleucine and valine, and X is any amino acid) (reviewed in (42)). There are a few examples where glutamate, E in the SUMO tetrapeptide motif can be substituted with an aspartate, D (reviewed in (2, 43)). It is to be noted that SUMOylation does not always occur within this motif and many lysines on substrate proteins that do not conform to the consensus motif are SUMOylated. An example of this is PCNA which is SUMOylated on two lysine residues: one of which resides in the consensus SUMO tetrapeptide motif while the other lysine residue does not and is present in a different location.

At the same time, presence of the SUMO consensus tetrapeptide does not guarantee attachment of Smt3p which would suggest that other factors such as subcellular localization and availability of target lysines may also contribute to modification by SUMO (42). Whilst the requirements for SUMOylation are quite simple, identification of proteins as SUMO substrates based on the presence of the ψKXE sequence is not sufficient. As a result, methods that evaluate SUMOylation of the global proteome are increasingly gaining favor (44–48).

SUMO Processing and DeSUMOylation by SUMO Proteases

SUMO proteins are synthesized in the cell as precursor proteins that need to be proteolytically cleaved to generate the mature C-terminus. SUMO proteases,

designated Sentrin/SUMO-specific proteases (SENPs) expose the terminal glycine residue that is linked to the target lysine on substrate proteins. The terminal glycine forms part of the diglycine motif – two glycine residues in tandem that are highly conserved in Ub and UBL family of proteins (15, 49, 50). The removal of conjugated SUMO from substrates is also catalyzed by the action of SUMO proteases making SUMOylation a cyclical and reversible process (reviewed in (14)). SUMO proteases remove SUMO from substrates in response to stimuli and in doing so generate a free pool of SUMO protein that can be conjugated to substrate proteins as needed. Two SUMO proteases have been described in yeast, Ulp1 and Ulp2 with the former essential for viability (reviewed in (14)). Deletions in Ulp2 in yeast result in cells that viable but grow abnormally and are hypersensitive to DNA-damaging drugs (51, 52).

It is of interest to study the mechanisms by which Ulp1 or SENPs proteases deSUMOylate proteins that are SUMO-modified because of their function in quick termination of SUMO-modification in response to various stimuli. The fewer number of proteases identified in the SUMO system (as compared to the much more complex Ub system) would suggest that these few proteases are highly active enzymes and/or there are SUMO proteases that may be identified in future studies.

Regulation of SUMOylation as a PTM

Control of Global SUMOylation

The primary method of regulating global SUMOylation in the cell is by regulating the expression or activity of the enzymes of the SUMOylation pathway. An overall increase in SUMOylation levels has been observed in response to ethanol stress, hydrogen peroxide and heat shock (47, 53). On the other hand, downregulation of the total SUMOylation is observed when the amount of reactive oxygen species is decreased (54). This type of control is exercised by formation of a disulfide bond between the active site cysteine residues of E1 and E2 enzymes rendering them defunct to participate in SUMOylation. In the Ub system, E1 enzymes, in addition to their primary role in activating Ub, can also establish substrate specificity by matching up with cognate E2s. The E2 enzyme Ubc9 itself possesses not only the capacity to catalyze SUMO transfer to the substrate, but also the capacity to select the lysine residue for modification. Because the SUMOylation pathway relies on a single E1 and E2, the regulation of these enzymes allows the cell to regulate the entire SUMO proteome at any given time. Stability of the E1 and E2 enzymes is another mechanism by which the cell can modulate the activities associated with the SUMO pathway. For example, chicken adenovirus GAM1 controls E1 turnover by binding to the SAE1 (Aos1p) component and recruiting Cullin-RING Ub ligases that target it for degradation by the proteasome (55, 56). Decrease in SAE1 levels causes a drop in SAE2 (Uba2p) levels concomitantly (56). GAM1 also reduces E2 enzyme Ubc9p levels to lower overall SUMOylation although the

mechanism of this inhibition is not fully understood (55). Pathogens like the bacterium *Listeria monocytogenes* infiltrate host defenses by impairing global SUMOylation by targeting Ubc9 as well as other SUMOylated proteins for proteasomal-mediated degradation (57).

Other methods to regulate SUMO pathway enzymes include sequestering of SUMOylation enzymes in different locations inside the cell. In budding yeast, the SUMO E3 ligase Siz1 is responsible for SUMOylation of septins. Siz1 is contained in the nucleus during interphase and at the onset of mitosis, it is release to the cytoplasm where it accumulates at septin rings and participates in their SUMOylation (58). Alcohol-induced stress in the same organism results in a rapid accumulation of the Smt3p protease Ulp1 in the nucleolus which consequently results in a dramatic increase in the levels and SUMOylation state of protein substrates (59). It is generally thought that SUMO E3 ligases (in addition to the identification of substrates by Ubc9) regulate SUMOylation as a PTM by displaying substrate preferences and in some cases, SUMO isoform-specific conjugation of particular targets (60).

Crosstalk with other PTMs

PTM modification by SUMO, Ub, and acetyl groups use lysine residues for attachment and it has become apparent over the course of several years that these PTMs can form a combinatory molecular switch to synergize or antagonize with each other adding yet another layer to protein regulation (reviewed in (61, 62)). As

these PTMs utilize lysines for covalent attachment, it is plausible that they compete with each other to bring about a more complex mechanism of protein regulation. A distinct function of Smt3p is to protect target proteins from Ub-mediated degradation. This is best exemplified in IκBα, an inhibitor of NF-κB transcription factor, which is degraded by Ub marking in cells responding to inflammation. In unstimulated cells, IκBα is protected from Ub-mediated degradation through SUMOylation which uses up the lysines otherwise targeted by Ub (63). Another example is cAMP response element binding protein (CREB) that is degraded in mammalian cells suffering from hypoxic stress. Overexpression of SUMO-1 stabilizes CREB in hypoxia which suggests a similar mechanism of protection offered by SUMOylation against ubiquitylation (reviewed in (64)).

There are cases, however, when Ub and SUMO act less antagonistically providing evidence for crosstalk between the two PTMs. Many substrates of SUMO contain a specific protein motif, denoted SUMO Interacting Motif (SIM) that is characterized by a sequence motif of hydrophobic amino acids V/I-X-V/I-V/I. These hydrophobic residues of the SIM domain contact a hydrophobic patch within SUMO (reviewed in (2)). RNF4, which belongs to a class of proteins referred as SUMO-targeted Ub ligase (STUBl), houses four SIMs which recognize SUMO moieties on target proteins. Upon recognition of polySUMO chains on proteins, RNF4 will start ligating Ub on the SUMO-modified protein thereby marking the protein for degradation by the proteasome (65).

In addition to the SUMO consensus tetrapeptide motif, the highly conserved phosphorylation-dependent SUMOylation motif (PDSM), that comprises the SUMO consensus motif and an adjacent proline-directed phosphorylation site (ψKXEXXSp; where serine is phosphorylated), has been described (reviewed in (42)). Some examples of proteins that are regulated both by SUMOylation and phosphorylation (referred as phospho-SUMOyl switch) are heat-shock factors (HSF), and peroxisome proliferator-activated receptor γ (PPARγ) and myocyte enhancer factor 2 (MEF2) (reviewed in (66)). In these cases, phosphorylation of the serine residue of the PDSM enhances SUMOylation as the negative phosphate group interacts with a basic path on the E2 Ubc9 enzyme (reviewed in (2)). Apart from such a phospho-SUMOyl switch, phosphorylation of SUMO substrates is also shown to regulate the SUMOylation state of the proteins. In some cases, dephosphorylation event at a PDSM causes a loss of the SUMO from the same PDSM motif. This decreased SUMOylation is coupled with increased acetylation indicating yet another variant of the SUMO consensus tetrapeptide – an acetyl-SUMOyl switch based on communication between SUMOylation and acetylation (67, 68). Acetylation of the target lysine within the PDSM blocks SUMOylation, leading to transcriptional activation through inhibition from SUMO-mediated repression (reviewed in (42)). SUMO proteins can be modified by acetyl groups further contributing to the regulation of SUMOylation. In mammals, SUMO-1 is targeted for acetylation which results in neutralizing of basic charges in its SIM docking sites and consequently

lowers its binding affinity to SIM domains in substrates such as PML, Daxx and PIAS proteins (67).

Biological Functions of SUMOylation

SUMOylation and Inhibition of Gene Transcription

Many SUMO-modified proteins reside in the nucleus but cytosolic targets have also been identified. The most prominent group of proteins that have been identified as SUMO substrates are transcription factors. These proteins are active in the nucleus where they regulate gene expression by modulating transcription. Mutations that prevent modification by SUMO in the transcription factors Elk-1 and Sp-3 result in an increase in transcription from respective promoters linking SUMO to repression of gene expression (69, 70). The role of SUMOylation in downregulating transcription is further reinforced by the observation that simply targeting SUMO or Ubc9 to promoters has been shown to reduce promoter activity (71). While not much is known about the general mechanism by which SUMOylation regulates transcription, it is becoming increasingly apparent that SUMO-modification results in interactions between transcriptional factors and transcriptional co-repressors such as chromatin modifying proteins that induce a more heterochromatic state (reviewed in (42)). For example, the histone deacetylase HDAC6 is able to bind to its co-repressor p300 only when it is SUMOylated ((1)). Another such interaction is seen in the case of PIAS proteins

(SUMO ligases) and Sp3 where PIAS binds strongly to SUMO-modified Sp3 than unmodified Sp3 (reviewed in (1)).

SUMOylation is Required for Maintenance of Genome Integrity

SUMOylation was first linked to DNA repair with the finding that thymine DNA glycosylase (TDG), a base excision repair enzyme that removes thymine or uracil from T-G or U-G mismatched base pairs, is a substrate of SUMO. TDG removes harmful DNA lesions leaving an abasic site to which it binds tightly. TDG has to be SUMOylated for it to relax its hold on the DNA helix which allows repair to occur quickly by downstream enzymes and frees up the TDG to seek other base lesions (72–74).

SUMO E3 ligases such as the PIAS proteins in mammals have been implicated in genome stability (75, 76). PIAS proteins are components of the DNA damage repair pathways and aid in the SUMOylation and subsequent recruitment of repair proteins to sites of double strand breaks (DSBs) (75). Although SUMOylation is important for response to DNA damage, it has also been demonstrated that deSUMOylation via the action of SUMO proteases also plays a role in genomic maintenance. The SUMO protease Ulp1 in yeast is an essential gene required for cell viability and Ulp2 mutant cells are viable but suffer from hypersensitivity to DNA damaging agents (51, 52). Taken together, these observations would suggest that SUMOylation and deSUMOylation may be coupled together in response to genotoxic

stress to regulate the assembly and disassembly of protein factors involved in the repair response.

SUMOylation in Chromosome and Sub-nuclear Structure

It is not surprising that a growing body of work links SUMOylation to the control of chromosome dynamics. In fact, all SUMO-pathway components (E1, E2, E3 and SUMO proteases) have shown genetic associations with maintenance of chromosome structure and segregation events. In *S. pombe*, complete disruption of the Smt3p homolog, Pmt3, results in cells that are barely viable and suffer from defects in mitosis and chromosome segregation (77). In the budding yeast *S. cerevisiae*, mutants of the E1-activating enzyme Uba2 display hypersensitivity to microtubule destabilizing drugs and arrest early in mitosis with short and frequently misaligned spindles (78). Yeast Ubc9 mutants are unable to appropriately activate the anaphase promoting complex (APC) that helps in the progression through mitosis (79).

SUMOylation has a well-defined role in dictating sub-nuclear architecture. This is best exemplified in the case of promyelocytic leukemia (PML) nuclear-bodies (PML NBs), where SUMOylation is necessary for maturation of the tumor suppressor protein PML which is the primary component of PML NBs (80, 81). PML NBs exemplify the function of the SUMO Interacting Motif (SIM) that are characterized by a tetrapeptide sequence motif of hydrophobic amino acids that make contact with a hydrophobic patch within the SUMO protein (reviewed in (2)).

SUMOylation of PML plays a critical role in its activity of recruiting partners, many of which are SUMOylated themselves. The ever increasing list of proteins of the PML complex also includes proteins that specifically recognize SUMOylated partners via their SIM domains. A combination of SUMO conjugation on target proteins and SIM domains on interacting partners is then essential for the formation of PML NBs (80, 81).

SUMO in Developmental Processes

Studies in several model systems have demonstrated that SUMOylation is critical developmental programs. In mitosis, specific roles for SUMO in regulating the normal execution of the cell cycle have been described since SUMO was first discovered two decades ago. In *S. cerevisiae* mutations in Uba2p or Ubc9p cause cells to arrest at the G2/M boundary indicating that SUMOylation is necessary for progression through the cell cycle (32, 79). Severe chromosome segregation defects are observed in mice lacking Ubc9 (81) and *S. pombe* mutants lacking Aos1p homologue (38). Several important targets of SUMOylation in mitosis have been identified including the cohesin and condensin complexes, components of which are known to be SUMOylated further linking SUMOylation and mitosis (reviewed in (82)). Another critical substrate of SUMOylation during mitosis is DNA topoisomerase II (TopII) where mutants that lack SUMOylation are unable to undergo normal chromatid dynamics (reviewed in (36)). In yeast, SUMO E3 ligase mutants do not display strong mitotic phenotypes (reviewed in (82)). This may be

because E3 ligase are redundant in their function for key cellular protein, or it is possible that Ubc9 is able to catalyze SUMO transfer during mitosis. Such regulation of the cell cycle is not limited to the E1, E2 and E3 enzymes alone – the SUMO protease Ulp1 has also been identified to be necessary for passage through the cell cycle. Ulp1 mutants in budding yeast are delayed at the G2/M boundary and suffer from increased chromosome mis-segregation (51). Deletion of Ulp2 is not lethal but cells lacking Ulp2 lose chromosome and are hypersensitive to microtubule destabilizing drugs (78).

During meiosis in yeast, SUMO is predominantly localized to synaptonemal complex (SC) and Zip3 which is a SUMO E3 ligase is crucial to initiating formation of the SC formation (84, 85). SUMO-mediated regulation of SC formation is a fairly conserved function across sexually reproducing organisms, but is absent from *Tetrahymena*. In higher eukaryotes such as mice and humans, SUMO localizes to specific regions of the chromatin that change as meiosis progresses (86, 87) suggesting that SUMOylation dynamically regulates proteins substrates throughout meiosis. In *Drosophila*, SUMOylation has been shown to modulate several developmental processes from wing morphogenesis, patterning of the eggshell and embryo to neurogenesis and metamorphosis which heavily indicates that SUMOylation is an important effector modification (reviewed in (88)). Several other *in vivo* studies have highlighted the importance of SUMO and its pathway proteins, Ubc9, PIAS and SENPs in embryonic development in numerous model systems (reviewed in (89)).

The SUMO Enigma

At any given time, the amount of SUMO present in the form of SUMO adducts is lower than the free form of SUMO (90–93). In addition, SUMOylation appears to be a transient modification due to the action of SUMO proteases and as a result, SUMO-modified proteins do not persist in the cell for long periods of time (5, 94, 95). These challenges make it difficult to investigate SUMO and its target proteins.

The most important experiment in studying the modification of a protein by SUMO is in the identification of the site of attachment of SUMO protein on the target protein. While mutating the lysine residue that forms the isopeptide bond is a good test for confirming SUMOylation state, it may negatively affect other PTMs that may utilize the same residue. Gene knockouts, conditional mutants, overexpression and dominant-negative studies may prove useful in studying SUMOylation but appropriate controls have to be utilized.

Herein the use of mass spectrometry-based proteomics approach that do not perturb the attachment sites of SUMOylation are extremely powerful and useful in the identification and characterization of SUMO substrates that exist as a smaller set of the total protein pool. An added advantage of these methods is the ability to identify several proteins at one time that give more insight into the SUMOylation state of the cell.

Tetrahymena is a Model System for studying SUMOylation

Ciliates are unicellular organisms that are characterized by presence of cilia on their cell surfaces and the presence of two nuclei unlike other eukaryotes. Each ciliate cell maintains both a transcriptionally inert germ-line micronucleus (mic) and a transcriptionally active somatic macronucleus (MAC). In the sexual life cycle, also referred as conjugation, two cells of different mating types pair together to initiate mating. In the early part of conjugation, the diploid mic in each partner will undergo meiosis to form haploid gamete nuclei. A single gametic nucleus will be transferred to another cell where it fuses with another gamete nucleus and form a zygotic nucleus. In the later stages of conjugation, broadly classified as MAC Differentiation, new progeny mics and MACs are derived from the zygotic nucleus (Figure 2) with the parental MAC degraded concomitantly. The new MACs (also termed anlagen) are derived from DNA amplification of the mic genome followed by genomic remodeling (reviewed in (96, 97)). The first type of type of genomic remodeling is chromosome breakage, in which the 5 chromosomes of the mic undergo ~50 breakages to form the ~250-300 acentric polygenomic chromosomes of the MAC. Chromosome breakage is a highly precise process with fragmentation of the chromosomes into smaller pieces and *de novo* addition of telomeres to newly broken ends (98, 99). The second type of genomic remodeling results in the elimination of mic-limited sequences from the anlagen. These sequences called internal eliminated sequences (IES) are removed by double-strand breaks in the DNA with the adjacent MAC-destined DNA being ligated back together (reviewed in

(100, 101)). It is now known that this type of DNA elimination in ciliates involves the same process of RNA-guided heterochromatin formation required for silencing of transposon elements in other eukaryotes (reviewed in (102, 103)).

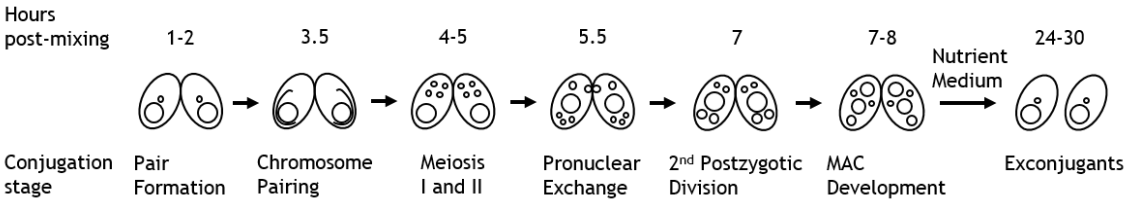

Figure 2. The sexual life cycle of *Tetrahymena*. In *Tetrahymena*, the diploid mic is transcriptionally silent and is responsible for germ-line functions. Conjugation (sexual reproduction) is a complex and dynamic process that starts cells pair (1-2 hours). Following pair formation, the nuclei elongate to form "crescent" structures (3.5 hours) and then undergo meiosis during which genetic material is exchanged between partner cells (4-5 hours). A single meiotic product is selected for pronuclear exchange (5.5 hours) which then undergoes two post-zygotic divisions to generate two new mic and MACs (7 hours). The process of MAC development (7-8 hours) involves extensive genomic remodeling and DNA amplification. When cells are returned to nutrient medium, they divide to generate daughter progeny (24-30 hours).

Tetrahymena is a useful model system for studying SUMOylation as a developmentally regulated process because it exhibits nuclear dimorphism – there is strong demarcation between somatic (MAC) and germ-line (mic) functions which makes it easy to study the vegetative and sexual life cycles separately. Mechanisms present in higher eukaryotes such as humans or other mammals that are either absent in other eukaryotic microbial model systems, or not as readily accessible in them as in *Tetrahymena*, are especially relevant. For example, lethal mutations can be harbored in the mic of a cell with a wild-type MAC that would presumably be

lethal in other model systems. The MAC genome has been sequenced and annotated and is comparable to that of multicellular organisms (104–106). Switching between the vegetative and sexual life cycles is as easy as mixing nutrient-starved cells of different mating types together. More importantly, the precise and tight regulation of nuclear processes that occur during conjugation provide a platform to study the various biochemical pathways that come into play during these critical processes. During MAC development the synchronized series of events in the new anlagen which include programmed DNA rearrangements in the form of site-specific DNA deletions (100, 106), *de novo* histone methylation (101), histone acetylation (107) and amplification of the MAC genome (100, 108) provide a dramatic and exaggerated series of biochemical events important in the nucleus. Genetic manipulations in *Tetrahymena* offer a natural complement to phenotypical studies for developing an overview of a pathway, and for identifying novel factors. A major challenge in current approaches to analyzing SUMOylation as a regulatory PTM is the low level of SUMOylated proteins (see *The SUMO Enigma* section above). The reasons for developing such a system in *Tetrahymena* are obvious - there are many model systems available for investigating the role of SUMOylation, however, few of them are amenable to studying the role of this PTM as they lack distinct developmental stages. An examination of the dynamic SUMOylation of substrates during these stages will reveal the role of low-level participating proteins. In this way, it will be possible to build functional links between the numerous substrates of Smt3p and the various developmental processes. We are interested in elucidating

the mechanism by which SUMOylation regulates programmed DNA rearrangements in the form of IES excision. The removal of these elements occurs in quite the dramatic fashion with approximately 6000 IES elements being excised followed by repair of resulting double strand breaks. This exaggerated phenomenon of IES excision occurs naturally as part of sexual life cycle of *Tetrahymena* and therefore provides a tremendous opportunity in studying the regulation of this process by SUMOylation. There are very few published studies that document the broad spectrum of proteins, involved in critical cellular processes, that are SUMOylated in response to different life cycles and herein *Tetrahymena* proves its versatility as a facile genetic model system.

CHAPTER 2 – SUMOYLATION IS DEVELOPMENTALLY REGULATED AND REQUIRED FOR CELL PAIRING DURING CONJUGATION IN *TETRAHYMENA*

Genetic analysis has revealed essential roles for *SMT3* in the survival and development of organisms ranging in complexity from lower eukaryotes such as yeast to mammals. Studies of the loss of SUMOylation components function reflect the impact of SUMOylation on entire organisms. Mutants for *SMT3 (SUMO), UBA2* as well as *UBC9* have been reported for most developmental models from yeast and *C. elegans* to mouse. Although different phenotypes have been described for different species, these studies have demonstrated that components of the SUMOylation pathway are essential genes in these organisms. In *S. cerevisiae*, there is an absolute requirement for *SMT3* in cell viability (33, 109). Deletion of SUMO pathway genes in the nematode *C. elegans* results in embryonic lethality, while downregulating the expression of SUMO conjugating enzyme Ubc9 in *Drosophila*, zebrafish, and mouse results in serious developmental defects (48, 81, 110).

The *Tetrahymena* genome encodes a single form of SUMO (referred to as *SMT3*; Table 1) (111), which shares 49% and 48% sequence identity with human SUMO-1 and SUMO-2, respectively. *SMT3* appears to have diverse roles inside the cell including, but not limited to, regulation of transcription, protein translation, DNA repair response, mitosis and meiosis etc. While *SMT3* is present throughout the

two life cycles of *Tetrahymena*, cells undergoing MAC development during conjugation contain particularly high concentrations of Smt3p (112), suggesting that SUMOylation may play particularly critical roles at this stage of *Tetrahymena* development.

Materials and Methods

Strains and Culture Conditions

Wild-type B2086 and CU428 strains of *Tetrahymena thermophila* were obtained from the *Tetrahymena* Stock Center, Cornell University, Ithaca, New York, USA. Strain B2086 contains 6-methly purine sensitive (*MPR1*) wild-type MACs with expression of mating type II. Strain CU428 contains 6-methyl purine sensitive (*MPR1*) wild-type MACs with expression of mating VII but contain micronuclei that have homozygous 6-methyl purine resistance (*mpr1-1/mpr1-1).* Wild-type and all other *Tetrahymena* strains were cultured in 1×SPP medium (2% proteose peptone, 0.1% yeast extract, 0.2% glucose and 0.003% $FeCl_3$) at 30 °C with shaking at 110 rpm for vegetative growth. Smaller volumes (1-10 mL) did not require shaking.

Mating Tetrahymena

For conjugation, *T. thermophila* strains were placed at 30 °C in 1×SPP and allowed to grow to logarithmic phase with an $O.D._{540}$ of 0.3 corresponding to 2×10^5 cells/mL. Cells were then washed with Starvation Buffer (10 mM Tris, pH 7.5) and then subsequently starved (16-24 hours at 30 °C) in 10 mM Tris (pH 7.5). Equal numbers of cells of different mating types were mixed together to initiate mating.

Construction of SMT3 and UBA2 germ-line knockout

The targeting construct consisted of a *neo3* cassette conferring paromomycin (pm) resistance placed under the control of the metallothionein (*MTT1*) promoter

(113). The drug resistance gene is flanked by sequences upstream and downstream of the *SMT3* (5' flank sequence was 1141 bp; 187801 to 188942 bp of scaffold 8254555 and 3' flank sequence was 1292 bp; 190002 to 191294 bp of scaffold 8254555) or *UBA2* coding sequence (5' flank sequence was 1004 bp; 472623 to 473627 bp of scaffold 8254719 and 3' flank sequence was 1148 bp; 476286 to 477434 bp of scaffold 8254719) (primers listed in Appendix Table 6). Wild-type B2086 and CU428 strains were then mated and the targeting construct was introduced at 2.5-3.5 hours post-mixing as described (114). Potential *SMT3* and *UBA2* micronuclear knock-out (KO) strains were identified by selection with paromomycin (confirming insertion of the *neo3* cassette) and 6-methyl purine (6mp; confirming successful conjugation and formation of new macronuclei). The heterozygous mic KO strains were further analyzed by PCR and genetic crosses to test strains (e.g. CU427) to confirm germ-line segregation of paromomycin resistance. These heterozygous germ-line transformants were crossed with "star" strains B*VI and B*VII that are deficient in donating a functional micronucleus. The subsequent uniparental transfer that occurs resulted in the generation of homozygous germ-line knockout heterokaryon strains (115). The micronuclei (germ-line) in these cells have homozygous deletions of the targeted gene, but the MAC genome is wild-type because the two partners of "star" crosses do not proceed through conjugation to form new MACs. PCR analysis as well as genetic crosses confirmed the genotype of each KO heterokaryon (see Table 2 and Figure 6).

Viability Test

To test the viability of the progeny of *SMT3* or *UBA2* homozygous heterokaryons strains, starved *SMT3* or *UBA2* heterokaryons strains (~2×10^5 cells/mL) were mixed to initiate mating and mating efficiency was assessed at 2 hours. Mating pairs were then isolated at 8-10 hours post-mixing in drops containing 1×SPP medium and placed at 30 °C. Cells in drops were examined at multiple times between 24 and 72 hours post-mixing. The number of cell divisions that occurred before death of homozygous *SMT3* and *UBA2* KO progeny was calculated by counting the total number of cells at the time of death, dividing by two (two exconjugants per pair) and calculating the cell divisions based on exponential growth. To determine successful completion of conjugation in *SMT3* or *UBA2* heterokaryons, cells were cultured in 1×SPP containing 80 μg/mL paromomycin. Progeny of control cell lines wild-type B2086.2 × CU428.1 were also tested for 6-methylpurine (7.4 μg/mL) resistance. Additional matings between *SMT3* and *UBA2* heterokaryons to wild-type partners were also performed to ensure generation of viable progeny (Table 2). To check for progression through conjugation, cells were fixed and stained with DNA-specific dye diamidinophenolindole (DAPI) as described below (see *Fluorescent and Confocal Microscopy*).

Creation of GFP-SMT3 and GFP-UBA2 constructs

To examine Smt3p and Uba2p localization in *Tetrahymena*, coding regions for Smt3p and Uba2p were amplified and cloned into the pENTR/D-TOPO plasmid

(Invitrogen, Grand Island, NY), which is used for recombination with the Gateway cloning system, to create pENTR-SMT3 and pENTR-UBA2 respectively. LR Clonase II (Invitrogen, Grand Island, NY) was used for directional cloning of *SMT3* or *UBA2* into the destination vector pBSmttGFPgtw containing an N-terminal *MTT1*-inducible GFP expression cassette cloned upstream of a cycloheximide-resistant Tetrahymena *rpl29* locus (described in (116)). For biolistic transformation, constructs were digested with HindIII or BlpI to produce linear plasmid with flanking *rpl29* sequences and transformed into starved *Tetrahymena* cells. Transformed cells were selected in SPP nutrient medium containing 12.5 μg/mL cycloheximide. To induce *GFP-SMT3* or *GFP-UBA2* expression, 0.1 μg/mL $CdCl_2$ was added to vegetative and mitotic cells. For mating cells, 0.05 μg/mL $CdCl_2$ was added when cells were first mixed and at 6 hours post-mixing. Cells were fixed as described below (see *Fluorescent and confocal microscopy*).

Fluorescent and Confocal Microscopy

GFP-SMT3 or *GFP-UBA2* expressing cells were fixed with 4% paraformaldehyde for 30 minutes at room temperature, then washed with 1×Tris-buffered Saline (TBS; 0.5 M Tris-Cl, 1.5 M NaCl, pH 7.6) for 10 minutes and stained with the DAPI at 1 μg/mL for 10 minutes. DAPI-stained cells were then placed on microscope slides and 5 μL of VectaShield Fluorescence Mounting Medium was applied to the cells. Fluorescent microscopy of vegetative strains expressing GFP transgenes (Figure 1) was performed using a Zeiss LSM 710 confocal microscope.

Digital images were processed using Zen 2009 (Carl Zeiss, Thornwood, NY) and Adobe Photoshop (Adobe Systems, Inc., San Jose, CA). Fluorescence images of mated cells (Figure 3) were obtained with an Olympus BX51TF model microscope and a 40X objective oil lens (UIS2/BFP1; Olympus).

Generating Conditional Mutants of SMT3 and UBA2

UBA2 knockout heterokaryons of different mating types were mixed with each other to initiate mating. Cells undergoing MAC development (corresponding to 8 hours post-mixing) were biolistically transformed using the *MTT1* promoter expressed *GFP-UBA2* transgene inserted at the *rpl29* locus (cycloheximide resistance). Progeny that were successfully transformed (paromomycin resistant and cycloheximide resistant) were complete gene knockouts for wild-type *UBA2* and expressed only the *GFP-UBA2* form. *SMT3* heterokaryons that were mated died in the first 24 hours after conjugation and we were unable to rescue cells using *GFP-SMT3* as described above. To generate a conditional mutant of *SMT3*, CU522 and CU527 strains that are sensitive to the drug paclitaxel (taxol; txs) were transformed with MTTp driven *FLAG-His6-SMT3* construct (gift of Dr. Joshua Smith, Missouri State University) which was incorporated at the *BTU1* locus resulting in progeny that are taxol-resistant (method originally described in (117)). *FLAG-His6-SMT3* expressing strains were further transformed with the *neo3* construct used earlier to generate *SMT3* KO heterokaryons. Cells were cultured in increasing concentrations of taxol and paromomycin to assort for *FLAG-His6-SMT3* copies and the reduction of

wild-type *SMT3*. Complete assortment of the wild-type *SMT3* gene away from the *FLAG-His6-SMT3* copy was demonstrated by performing two-step Reverse Transcriptase-PCR (RT-PCR) on RNA obtained from *FLAG-His6-SMT3* expressing strains which showed amplification of the *FLAG-His6-SMT3* product only.

Assessment of Pair Formation

To assess pair formation, conditional mutants of *SMT3* or *UBA2* were cultured using three different conditions. The "+Cd" cultures were grown in SPP medium supplemented with 0.1 μg/mL $CdCl_2$ for 24 hours at 30 °C. These cells were then starved in Starvation Buffer containing 0.1 μg/mL $CdCl_2$ for 24 hours at 30 °C after which they were mixed to initiate mating pair formation. The "−Cd" cultures were grown and starved in the absence of $CdCl_2$ and then mated to assess pair formation. The "+Cd Addition" cultures, were treated as described for the "−Cd" set except that 2 hours prior to mixing, cells were supplemented with 0.1 μg/mL $CdCl_2$ and then mated. The percentage of cells forming pairs was calculated as the number of cells in pairs divided by the total number of cells in the sample (paired plus single cells).

Preparation of Whole Cell Extracts

100 mL of vegetative wild-type strains B2086.2 and CU428.1 at log-phase (2×10^5 cells/mL) and mated cells (at 0, 2, 5, 7, 10, 24, and 30 hours post-mixing,) were harvested by centrifugation, washed twice with Starvation Buffer (10 mM Tris-

Cl, pH 7.5) and then resuspended in 2 mL of lysis buffer (8 M Urea, 50 mM Tris, pH 8.0) containing 50 mM N-ethylmaleimide and lysed by sonication. The resulting lysate was centrifuged at 4 °C for 30 min at 77,000×*g* in a SW41 rotor (Beckman Coulter). The clarified supernatant was then prepared for Western Blotting as described below.

Western Blotting and Coomassie Staining Analyses

Clarified lysates from wild-type strains were prepared by the addition of SDS loading buffer to a final concentration of 1× (50 mM Tris, pH 6.8, 10% glycerol, 2% SDS, 100 mM DTT, and 0.05% bromophenol blue). Samples were separated on 8% Bis-Tris Gels using Tris Glycine SDS running buffer and then transferred to PVDF membrane (Pall Corporation, Port Washington, NY) at 100 V for 2 h. Subsequent Western blot analysis was then performed with anti-Smt3p and horseradish peroxidase-conjugated secondary antibodies (Jackson Immunoresearch, West Grove, PA). Smt3p antibodies were custom generated by ProteinTech Group Inc. (Chicago, IL). Rabbits were immunized with a peptide corresponding to amino acids 26-43 of *Tetrahymena* Smt3p (FFKIKKTTQFKKLMDAYC) and then antibodies were affinity purified from the resulting serum using the same peptide. Results were visualized with the ECL Prime chemiluminescence detection system (Amersham Biosciences, Piscataway, NJ). Coomassie staining was performed by first treating the SDS gel with fixing solution (10% glacial acetic acid and 25% isopropanol) for 3 hours and then incubating it in Coomassie R250 staining solution (0.05% Coomassie

R250, 10% acetic acid and 25% methanol) overnight at room temperature. Destaining was done in 10% glacial acetic acid for 3 hours.

Results

Tetrahymena Smt3p and Uba2p primarily accumulate in the macronucleus

Previous studies on *Paramecium tetraurelia* revealed that RNAi induced silencing of *UBA2* or *SUMO* resulted in the failure of programmed DNA rearrangements (111), but the underlying mechanism responsible for this effect was not established. *Tetrahymena* was selected for additional studies because of technical advantages (gene knockouts, higher mating efficiency) and the opportunity for comparative analysis with *Paramecium*. To initiate our analysis of SUMOylation we searched the *Tetrahymena thermophila* genome for homologs and identified a single gene encoding SUMO (118) with reciprocal top BLAST hits to SUMO proteins in *S. cerevisiae, Drosophila* and human. We named the *Tetrahymena* gene *SMT3* (TTHERM_00410130) (111), consistent with *S. cerevisiae* nomenclature (Suppressor of Mif Two 3) (Table 1). The alignment of *Tetrahymena* Smt3p with human SUMO isoforms, *S. cerevisiae*, *S. pombe* and other protozoa is shown in Figure 3A. *Tetrahymena* Smt3p shares 50% identity with *S. cerevisiae*. The N-terminal regions show substantial divergence, but most of the protein is highly conserved including the diglycine motif conserved in the C-terminus of most UBLs. A *Tetrahymena* homolog for *UBA2* was previously reported based on its identity with *Paramecium tetraurelia* (TTHERM_00391590) (111). That study evaluated the developmental expression of transcripts from the *Tetrahymena UBA2* and *SMT3* (SUMO) genes with northern hybridizations. Both showed substantial increases in expression during sexual reproduction, consistent with subsequent data from whole

genome microarrays (40). The developmental pattern of expression and sequence identity between these genes when compared with the previously examined *Paramecium UBA2* and *SUMO* genes led us to perform additional studies.

To evaluate whether the cellular location of *Tetrahymena* SUMOylation pathway proteins is consistent with *Paramecium tetraurelia* and other eukaryotes, we examined the localization of Smt3p and Uba2p. Localization of Smt3p and Uba2p during vegetative growth was examined by expressing *GFP-SMT3* or *GFP-UBA2* transgenes from a metallothionein promoter regulated by cadmium. Both Smt3p and Uba2p were observed in the somatic MAC but not the germ-line mic in vegetative *Tetrahymena* (Figure 3B). Based on the coding sequence of *SMT3*, we generated rabbit polyclonal anti-Smt3p antibodies by immunizing animals with residues 21-43 of the full-length *Tetrahymena* Smt3p protein and used them to examine localization of Smt3p in vegetative cells (described in Materials and Methods). As seen in Figure 3C the predominant signal is in the macronucleus and no signal can be detected in the micronucleus. The signal at the cell cortex was detected with secondary antibody alone (no primary antibody) and is therefore unlikely to be significant. Additional experiments with the GFP transgenes showed that increasing or decreasing cadmium concentrations changed the strength of the GFP signal but did not alter its nuclear localization (data not shown). Although a small fraction of Uba2p and Smt3p are likely present in the cytoplasm we are confident that the signal is predominantly macronuclear in vegetative cells. Studies of other developmental model systems such as mice (22, 23), *Drosophila* (24, 25)

and yeast (26) have reported that both Smt3p and Uba2p localize predominantly to the nucleus. This is expected as many of the proteins modified by SUMO are nuclear, including promoter-specific transcription factors, DNA repair proteins and chromatin-associated proteins. Our results in *Tetrahymena* are consistent with the role of SUMOylation in somatic nuclear processes such as the regulation of transcription, which is limited to the MAC in vegetative cells.

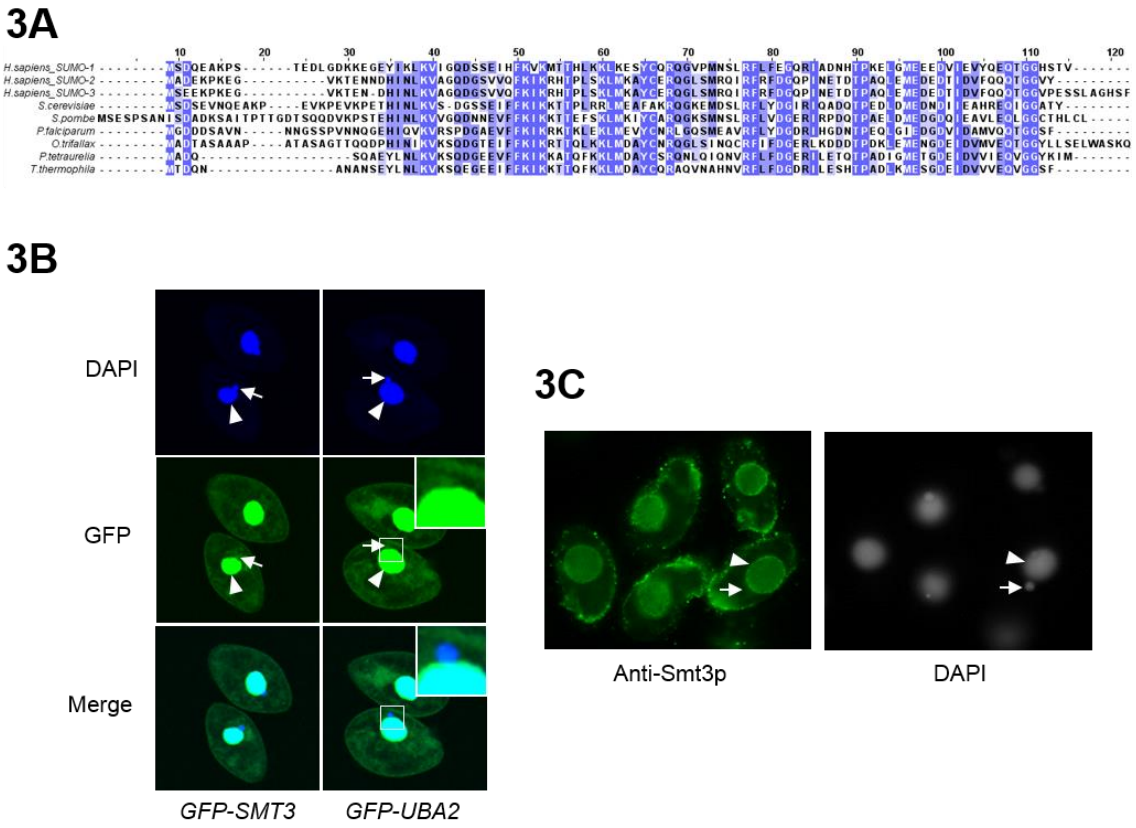

Figure 3. Smt3p localizes to the somatic MAC during vegetative growth. (A) Sequence alignment of Smt3p isoforms from 7 species. Dark shading represents residues with highest percent identity across all sequences. Light shading indicates less conserved residues. The sequences and the accession numbers of Smt3p are as follows: *H. sapiens* SUMO-1, NP_003343.1; *H. sapiens* SUMO-2, NP_008868.3; *H. sapiens* SUMO-3, NP_008867.2; *S. cerevisiae*, NP_010798.1; *S. pombe*, NP_596035.1; *P. falciparum*, PFE0285c in PlasmoDB; *O. trifallax*, Contig18025.0.g12 in OxytrichaDB; *P. tetraurelia*, GSPATG00013187001 in ParameciumDB; and *T. thermophila*, TTHERM_00410130 in *Tetrahymena* Genome Database. (B) Nuclear localization of Smt3p and Uba2p during vegetative growth. *GFP-SMT3* and *GFP-UBA2* constructs were transcribed from the metallothionein 1 promoter upon addition of cadmium. The GFP signal prominently localized in the somatic MAC but was absent from the germ-line mic (inset for the *GFP-UBA2* is shown). (C) Immunofluorescence of log-phase wild-type cells treated with *Tetrahymena* polyclonal anti-Smt3p antibodies show that the signal is predominantly macronuclear. Cortical staining was observed in pre-immune controls (data not shown). Arrow-heads indicate MAC and arrows indicate the position of the mic.

SUMOylation increases during conjugation in T. thermophila

Anti-Smt3p antibodies used in Figure 3C were also utilized to determine SUMO expression throughout conjugation by western blot analysis. Figure 4A shows schematic representations of the stages when protein was prepared for analysis. Figure 4B shows a typical SUMOylation pattern with several reactive bands consistent with a range of proteins conjugated to Smt3p. The image is a short exposure to emphasize the difference in signal between vegetative cells and conjugating cells, but longer exposure show a large number of bands in the vegetative protein samples. The arrow head labeled Smt3p indicates the expected migration of free Smt3p. Experiments with high percentage gels (20%) showed that Smt3p migrates with a mass of approximately 13.5 kDa, but on the 10% gel shown in Figure 4B proteins below 20 kDa are not well resolved. A slightly higher apparent molecular mass than the theoretical molecular mass (11.4 kDa) of Smt3p is not surprising as SUMO proteins have been reported to exhibit anomalously slow migration on SDS-PAGE gels (1). The high molecular mass bands observed at the 40-200 kDa range are Smt3p substrates that represent a diverse set of target proteins that are modified by Smt3p in vegetative and mating cells. In mating cells, we observed that Smt3p adduct formation increases as cells progress through conjugation with the highest signal observed between 7 and 10 hours post-mixing which corresponds to anlagen formation implicating SUMOylation in its role in MAC differentiation (Figure 4A and 4B). The bar graph in Figure 4C provides a quantitative assessment of the increased signal above 40 kDa relative to total

protein as measured by Coomassie staining. This increase in Smt3p signal is consistent with microarray expression data, which shows an increase in transcript levels during conjugation including anlagen formation (40). These results demonstrate that SUMOylation occurs differentially between vegetative and mating *Tetrahymena* with a peak in SUMO conjugates observed during MAC differentiation.

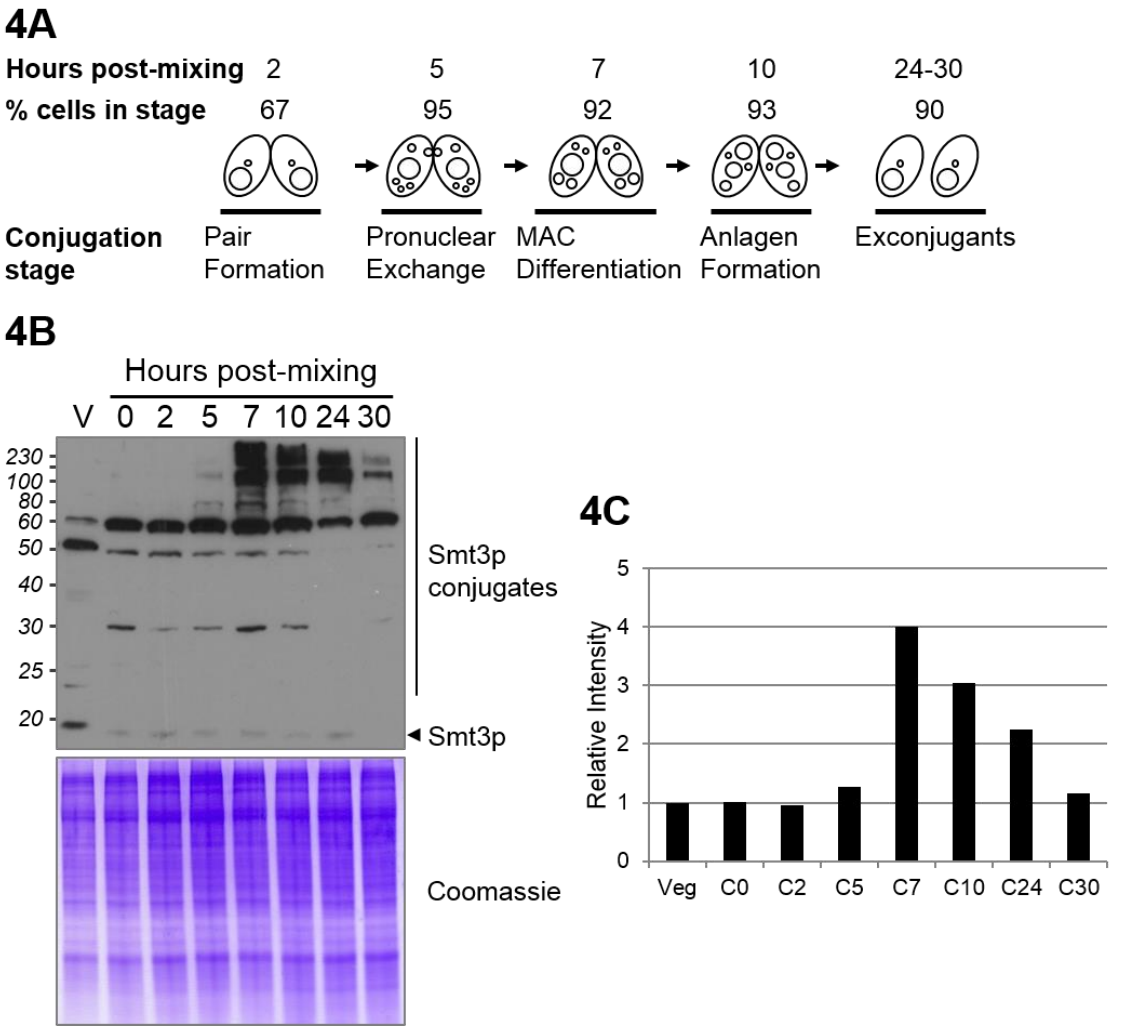

Figure 4. Differential modification of proteins by Smt3p during conjugation versus vegetative growth. (A) Schematic representation of the stages of conjugation at which cells were lysed. The following stages are shown: Pair Formation; Pronuclear Exchange; MAC Differentiation; Anlagen Formation; and Exconjugants. (B) Whole cell extracts (WCEs) were prepared from vegetative and mating B2086 and CU428 wild-type cells at 0, 2, 5, 7, 10, 24, and 30 hours post-mixing. WCEs were analyzed by 10% SDS-polyacrylamide gel electrophoresis followed by Coomassie blue staining (for loading control; bottom panel) and Western blot using polyclonal anti-Smt3p antibody (top panel). Arrow-head indicates the expected position of free Smt3p (~13 kDa) and high molecular weight proteins that are likely to be Smt3p conjugates. Coomassie stained gel is included as a control to evaluate equivalent sample loading. (C) The bar graph provides a quantitative assessment of the immunoblot signal across different time points after normalizing for sample loading (Coomassie stained gel).

To determine the cellular location of the increased Smt3p we performed immunofluorescence with the same anti-Smt3p antibody used in Figure 4. Mating cells were fixed at various times during conjugation leading up to the beginning of MAC development (9 hr). A low signal relative to background fluorescence masked specific localization in early stages of conjugation. As conjugation progressed, Smt3p was detected in parental MACs by the end of meiosis, and the antibody revealed a strong accumulation in the developing anlagen (Figure 5). We also used the GFP-Smt3p expressing strains (described in Figure 1) and found that Smt3p first localized to parental MACs early in conjugation and later accumulated in the developing anlagen which mimicked our results with the anti-Smt3p immunofluorescence analysis (Figure 5). Both techniques showed that the signal in the parental MAC disappeared as the anlagen developed and before the MAC was degraded. We did not detect Smt3p in the micronuclei or meiotic products during conjugation using anti-Smt3p or GFP-Smt3p. Although the GFP-Smt3p strains have the disadvantage of expression from the inducible *MTT1* promoter, the consistent results using antibodies and GFP fusions give us confidence in our observations.

In *Paramecium* GFP-Uba2p localized to the developing MAC (111). Antibodies against *Tetrahymena* Uba2p are not available for immunofluorescence, but GFP-Uba2p was expressed from the inducible *MTT1* promoter in the absence of wild-type Uba2p (described in Materials and Methods). As shown in Figure 3, GFP-Uba2p is located in the old MAC during the meiotic "crescent" phase (prophase I) of conjugation. The signal remains in the old MAC at later stages of meiosis and unlike

Smt3p, Uba2p is clearly visible in the meiotic products labeled “meiotic haploid products” in Figure 5. The signal is also detectable in haploid products during the period of pronuclear exchange (Figure 5). As the developing MAC (anlagen) appear they contain increasing Uba2p signal and the signal is simultaneously reduced in the parental (old) MAC. Consistent with our observations in vegetative cells, the signal is not detected in micronuclei at the two mic two MAC stage of development. Interestingly, *Paramecium* GFP-Uba2 expressed from its endogenous promoter showed little signal in the parental MAC and the primary signal appeared in the anlagen (111). The *Tetrahymena* Uba2p and Smt3p localization results (Figure 5) along with the immunoblot (Figure 4) are consistent with a major increase in SUMOylation of nuclear proteins during conjugation, particularly in the developing MAC. The absence of GFP-Smt3p localization in the meiotic products in contrast to the GFP-Uba2p localization to those structures is not readily explained, but it is clearly not an issue of protein abundance (the result was independent of the level of expression) or stability (since the signal in the MAC was robust).

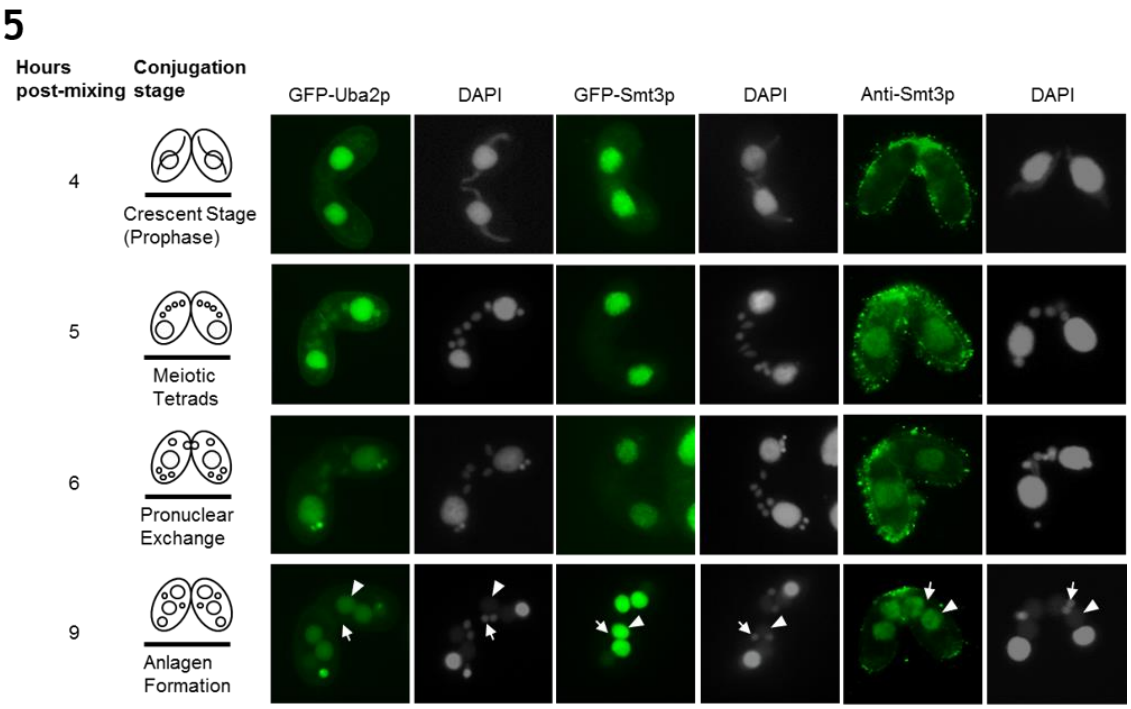

Figure 5. Localization of Uba2p and Smt3p in conjugation. *UBA2* deletion cells were rescued with the *MTT1*-driven *GFP-UBA2* construct to generate N-terminally tagged GFP-Uba2p expressing cells. Cells were mated and then fixed and stained with DAPI at various times throughout conjugation. GFP images were taken to visualize the distribution of *UBA2* in mating pairs. *GFP-SMT3* was transformed into B2086 and CU428 wild-type strains to produce N-terminal GFP-Smt3p tagged strains. The GFP-Smt3p expressing strains were mated and fixed at time-points shown in the figure. Smt3p localization was also visualized in mated wild-type cells using anti-Smt3p antibodies. A schematic of nuclear morphologies as observed in wild-type mating is drawn for reference. The following developmental stages are observed: "crescent" micronuclei (prophase meiosis I) (4 hr); meiotic haploid products generated after completion of meiosis II (5 hr); pronuclear differentiation and exchange (6 hr); second post-zygotic mitosis (7 hr); and macronuclear anlagen formation and nuclear processing (9 hr). Arrow-heads indicate selected anlagen, arrows indicate selected mic.

Complete Deletions of SMT3 and UBA2 Result in Cell Lethality

In budding yeast and invertebrates such as the nematode *Caenorhabditis elegans*, there is a single *SMT3* gene which is essential for viability (28, 29).

Deletions in the SUMO-activating enzyme *UBA2* or the SUMO-conjugating enzyme *UBC9* are lethal in the budding yeast (32, 33, 79), *C. elegans* (119), *A. thaliana* (120) and in mouse (81). In *Paramecium tetraurelia*, RNAi silencing of *UBA2* or *SUMO* had no detectable effect on vegetative cells but arrested conjugating cells. To test whether *SMT3* and *UBA2* are essential genes in *Tetrahymena*, we generated micronuclear (germ-line) deletion strains that were subsequently mated to produce complete (mic and MAC) deletions. The initial heterozygous mic knockout strains were selected with paromomycin after transformation with knockout (KO) constructs containing the neomycin resistance cassette (*neo3*) flanked by sequences upstream and downstream of each coding regions (Figure 6A). The resulting heterozygous knockout cells were cultured without paromomycin selection to allow phenotypic assortment of the drug resistant (KO) alleles and identification of paromomycin sensitive cells with fully wild-type MACs. These heterozygous knockout cells were mated with "star" strains containing defective micronuclei and wild-type MACs. These crosses result in abortive conjugation, in which paired cells complete meiosis and exchange genetic material, but separate without making new macronuclei. These post-conjugation cells emerge with homozygous micronuclei, but retain their original MACs genetically wild-type for *SMT3* and *UBA2* (Figure 6A and 6B) (115). Phenotypically these cell lines are paromomycin sensitive, but we used PCR amplification to identify the strains homozygous for the *neo3* cassette in place of the corresponding coding region in the micronucleus (Figure 6C). These homozygous heterokaryon knockout strains of *SMT3* or *UBA2* were then mated and

individual pairs were isolated to nutrient medium. Analysis of >100 pairs from each cross revealed that the majority of pairs failed to survive (Table 2). Pairs that did survive were tested for paromomycin-resistance to determine whether they were true exconjugant progeny or pairs that aborted conjugation. This test was used because the parental strains were phenotypically paromomycin-sensitive (MAC genotype) but their germ-line micronuclei were homozygous for paromomycin resistance. The formation of a new MAC would result in paromomycin resistance. As indicated in Table 2, none of the surviving lines were paromomycin resistant and most likely they were abortive mating pairs containing wild-type MACs. Control crosses of the same knockout heterokaryon strains with wild-type cells (B2086 or CU428) resulted in high survival (~93 %) of true exconjugant progeny that were paromomycin resistant. This demonstrates that the knockout heterokaryons contain fertile micronuclei. The data are consistent with a lethal phenotype for complete deletions of *SMT3* or *UBA2.* This has been reported in other organisms and we conclude that both are essential genes in *Tetrahymena*.

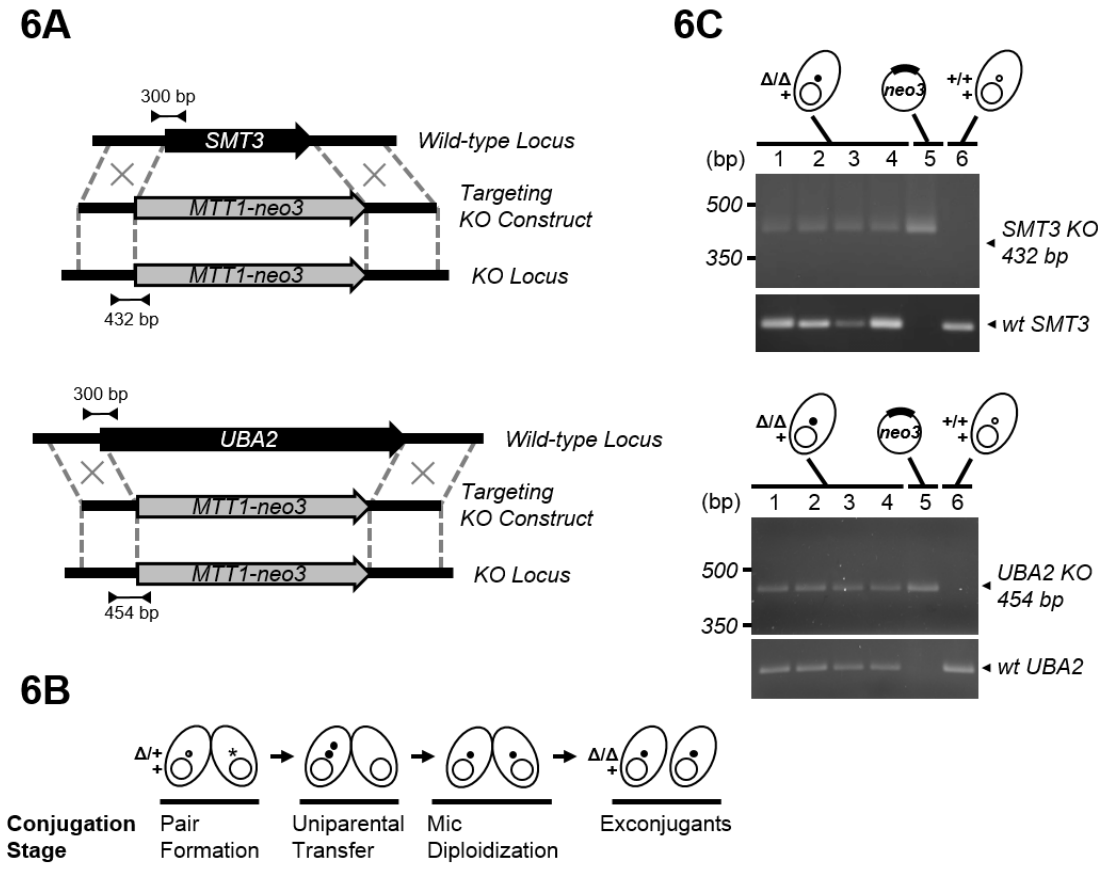

Figure 6. Generation of *SMT3* and *UBA2* knockout heterokaryons. (A) Schematic drawings of *SMT3* and *UBA2*, their targeted loci, the knockout construct as well as the resulting gene knockout formed by homologous recombination. The entire *SMT3* and *UBA2* coding sequences were replaced by insertion of the *neo3* knockout cassette that confers paromomycin resistance. Arrows indicate the locations of primers used for genotyping PCR from total DNA. (B) To generate *Tetrahymena* with homozygous deletions in the mic, a type of abortive mating called round I genomic exclusion was utilized wherein wild-type cells were crossed with 'star' strains that have defective micronuclei. Star strains are able to form mating pairs with wild-type cells but do not provide a migratory gametic micronucleus to the wild-type partner following meiosis. The single haploid micronucleus that is donated by the wild-type partner to the star partner, then undergoes endoreplication, to form a homozygous, diploid micronucleus in each conjugant. At this point, the mating pair aborts conjugation and separates. (C) Total DNA isolated from wild-type CU428 (lane 6), *SMT3* (lanes 1-4; top panel) and *UBA2* (lanes 1-4; bottom panel) knockout homozygous heterokaryons was PCR amplified with primers shown in (A). Positions of the bands for the disrupted (KO) loci as well as wt genes are indicated. Genotypes of *SMT3* and *UBA2* homozygous germ-line knockout heterokaryon strains and wild-type strains are illustrated at the top. Lane 5 in both gels contained the corresponding KO construct as template.

Interestingly, there was a substantial difference in the timing of post-conjugation death for *ΔSMT3* and *ΔUBA2* strains. *ΔSMT3* progeny died before the first post-zygotic cell division, but *ΔUBA2* progeny died much later, 6-8 cell divisions after mating (~72 hours post-isolation). Clearly *UBA2* expression is not required from the zygotic macronucleus (anlagen) to complete conjugation. The *SMT3* and *UBA2* knockout heterokaryon crosses were monitored by DAPI staining to detect any grossly aberrant nuclear events (i.e. meiosis, pronuclear fusion, mitotic divisions and anlagen development), but none were detected. The *ΔUBA2* vegetative progeny have the normal distribution of 1 MAC and 1 mic prior to death. The *ΔSMT3* progeny arrest in conjugation at the 2 MAC, 2 mic stage, but the events prior to that appear normal (data not shown). We also examined the *ΔSMT3* progeny and *ΔUBA2* progeny for defective programmed DNA elimination. Using a PCR strategy developed previously (116) we examined DNA from cells at the time of mixing (time 0) and 24 hours later from several hundred conjugating pairs collected after mixing the same culture. A total of five eliminated DNA elements were examined for each cross. The results for the M and R elements in *ΔSMT3* progeny are shown in Appendix Figure 15. For each eliminated element the product expected for accurate DNA processing was detected and there was no evidence for the inhibition of DNA elimination. The explanation for the shorter life span of *ΔSMT3* cells compared with *ΔUBA2* is not clear but could result from faster depletion of parentally expressed Smt3p.

Table 2. *SMT3* and *UBA2* are essential genes.

Types of mating cells[a]	**No. individual pairs examined**	**No. wells with more than 2 cells/drop**	**No. of clones completed conjugation**[b]
Δ*SMT3-neo3-1* × Δ*SMT3-neo3-2*	110	34	0
Δ*UBA2-neo3-1* × Δ*UBA2-neo3-2*	124	32	0
Δ*SMT3-neo3-1* × B2086	110	91	81
Δ*UBA2-neo3-1* × B2086	110	96	88
Δ*SMT3-neo3-2* × CU428	110	89	84
Δ*UBA2-neo3-2* × CU428	110	92	83
B2086 × CU428	88	82	82

[a]Each row represents data obtained in the course of three independent experiments.

[b]To distinguish between cells that completed conjugation (progeny cells) and cells that have aborted conjugation, wild-type B2086 and CU428 strains were tested for resistance with 6-methyl purine (only progeny cells should be resistant) and cells from the *ΔSMT3-neo3 × ΔSMT3-neo3* and *ΔUBA2-neo3 × ΔUBA2-neo3* mating were checked for paromomycin resistance (progeny should be paromomycin resistant).

Gene Rescue with MTT1 Promoter Converts the SMT3 or UBA2 Deletion Lines into Conditional Mutants

As *SMT3* and *UBA2* are essential genes, complete deletions resulted in cell death making it difficult to obtain information on the null phenotype of these genes. This problem was circumvented by the creation of conditional mutants of *SMT3* and *UBA2* in which the expression of these genes could be regulated by addition of $CdCl_2$ to culture medium. As *ΔUBA2* progeny survive for up to 72 hours post-mixing, the parental *UBA2* heterokaryons were mated *en masse* and at 8 hours post-mixing (corresponding to formation of new anlagen), mating cells were biolistically transformed with the *GFP-UBA2* construct (placed under *MTT1* control) to rescue

the lethal phenotype (Figure 7A). The only expressed functional copy of the *UBA2* gene in these cells was the *GFP-UBA2* transgene under *MTT1* control. For *ΔSMT3* mated cells, we were unable to rescue the lethal progeny with the *GFP-SMT3* transgene, possibly due to insufficient expression of the introduced copies at the end of conjugation. To bypass this issue, we generated *Tetrahymena* cell lines that were somatic transformants expressing introduced FLAG-His6 epitope tagged *SMT3* (also driven by the *MTT1* promoter) at the *BTU1* locus in CU522 and CU725 strains. Incorporation at the *BTU1* locus in these strains confers resistance against the drug taxol (paclitaxel) that are otherwise taxol-sensitive. Next, we transformed these *FLAG-His6-SMT3* expressing strains with the *neo3* knockout construct that replaces the endogenous *SMT3* gene and then cultured cells in increasing concentrations of paromomycin and taxol to select for cells in which wild-type copies of *SMT3* had been assorted out and only *FLAG-His6-SMT3* remained as the functional copy (Figure 7B and C). Reverse transcription PCR was used to confirm the absence of the wild-type SMT3 transcript (Figure 7C). To determine whether placing *SMT3* and *UBA2* under the control of the *MTT1* promoter would lead to a conditional, Cd-regulated mutant, we examined the growth rates of these cell lines compared to that of wild-type strains. The growth rate of the conditional strains is similar to that of wild-type cells when the culture medium is supplemented with cadmium, but significantly slower when these strains are cultured in cadmium-free medium (Figure 7A and B). The difference between the lethal phenotype of the complete gene deletion strains and the slow growth phenotype of the conditional cell lines in

the absence of cadmium can be explained by “leaky” expression of the *MTT1* promoter even in the absence of cadmium. This has been reported previously (113) and we have seen this in our own analysis of *MTT1* promoter constructs. These results demonstrate a conditional phenotype in the absence of $CdCl_2$ that is consistent with Smt3p and Uba2p depletion.

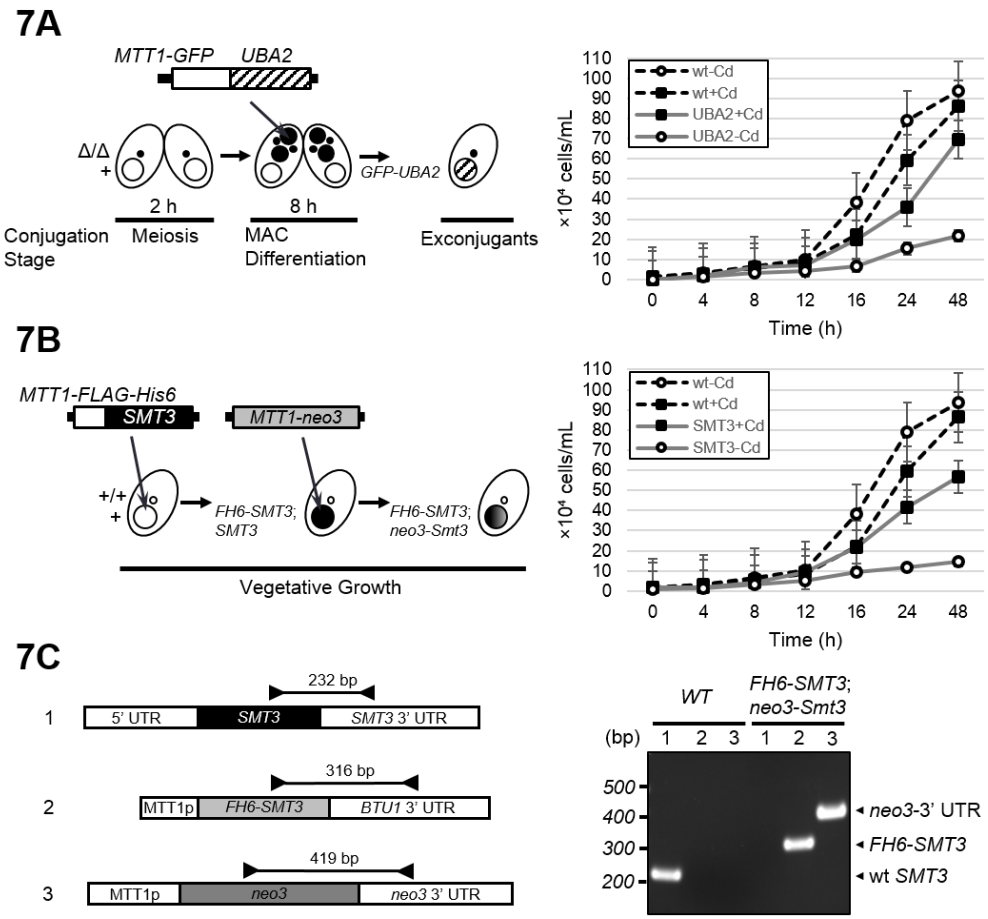

Figure 7. *Tetrahymena* expressing the *GFP-UBA2* or *FLAG-His6-SMT3* transgene regulated by the *MTT1* promoter behave as conditional mutants. (A) Schematic drawing of *ΔUBA2* homozygous heterokaryon strains that were mated and transformed with the *GFP-UBA2* transgene regulated by the *MTT1* promoter and inserted at the rpl29 locus. The resulting *UBA2* conditional cell lines and wild-type CU428 strains were grown vegetatively and transferred into 10 mL of SPP medium at an initial concentration of 200 cells/mL with either 0.1 μg/mL cadmium or no cadmium. At 0, 4, 8, 12, 20, 24, and 48 hours after placement in growth medium, cells were fixed and scored using a hemocytometer. (B) *ΔSMT3* homozygous homokaryon progeny were lethal within 24 hours. Schematic drawing shows the transformation of CU522 strain with *MTT1-FLAG-His6-SMT3* at the *BTU1* locus. The same strain was then transformed with the *neo3* knockout construct (*MTT1* promoter) disrupting the endogenous *SMT3* gene. These conditional *SMT3* cells were scored for growth as described in (A) above. (C) RT-PCR reactions to detect wt *SMT3* transcripts in conditional mutants. The locations of the primers (arrows) used to amplify regions from the cDNA are shown on the left. Lanes 1, 2, and 3, on the gel correspond to PCR reactions using the primer sets (1, 2, and 3) indicated on the diagrams to the left. No wt *SMT3* transcript is detected in the *SMT3* conditional mutant cell line.

SMT3 and UBA2 requirement in mating pair formation in conjugating cells

We know from previous studies in *Paramecium tetraurelia* that silencing *UBA2* results in conjugation arrest during MAC development (111). Using our conditional Smt3p and Uba2p cell lines we tested the effect of Smt3p or Uba2p depletion on conjugation in *Tetrahymena*. Conditional mutants of different mating types were cultured overnight in nutrient medium with or without cadmium to either promote or reduce expression of *SMT3* or *UBA2* transgenes respectively. Each culture was washed and cultured in starvation buffer for 24 hours under the same plus (+Cd) or minus cadmium (−Cd) conditions. Two hours before mating, 0.1 μg/mL of cadmium was added to a flask containing a portion of the cadmium free culture (referred to as "+Cd Addition" cells). Wild-type cells were carried through the same procedure under the same conditions. Cells from cadmium and non-cadmium treated cultures were mated separately. Cells were evaluated at 2 hours and 8 hours post-mixing for pair formation and cell samples were DAPI stained to examine the nuclear events during conjugation. Smt3p and Uba2p conditional mutant cultures that had not been exposed to cadmium were unable to form mating pairs (Figure 8). In contrast, the conditional mutants that had been exposed to cadmium had greater than 70% pair formation and progressed normally through the nuclear events of conjugation (data not shown). Smt3p and Uba2p conditional cells that were cultured without cadmium, but treated with cadmium two hours prior to mating were able to form pairs and progress through the normal nuclear events of conjugation as assayed by DAPI staining (Figure 16). The finding that conditional cells can recover their ability

to pair clearly demonstrates a dependence on Smt3p and Uba2p expression immediately prior to conjugation, and eliminates the possibility that cells do not pair as a result of defective nuclei or other damage that could occur during vegetative growth. Wild-type control cells, both cadmium and non-cadmium treated, showed greater than 80% pairing and normal progress through conjugation as expected (Figure 8). The non-pairing phenotype of Uba2p and Smt3p depleted cells was unexpected and prevents the use of these cell lines to examine later stages of conjugation. Nevertheless, the elevated expression of SUMO pathway proteins and increased SUMOylation during MAC development make additional critical functions likely.

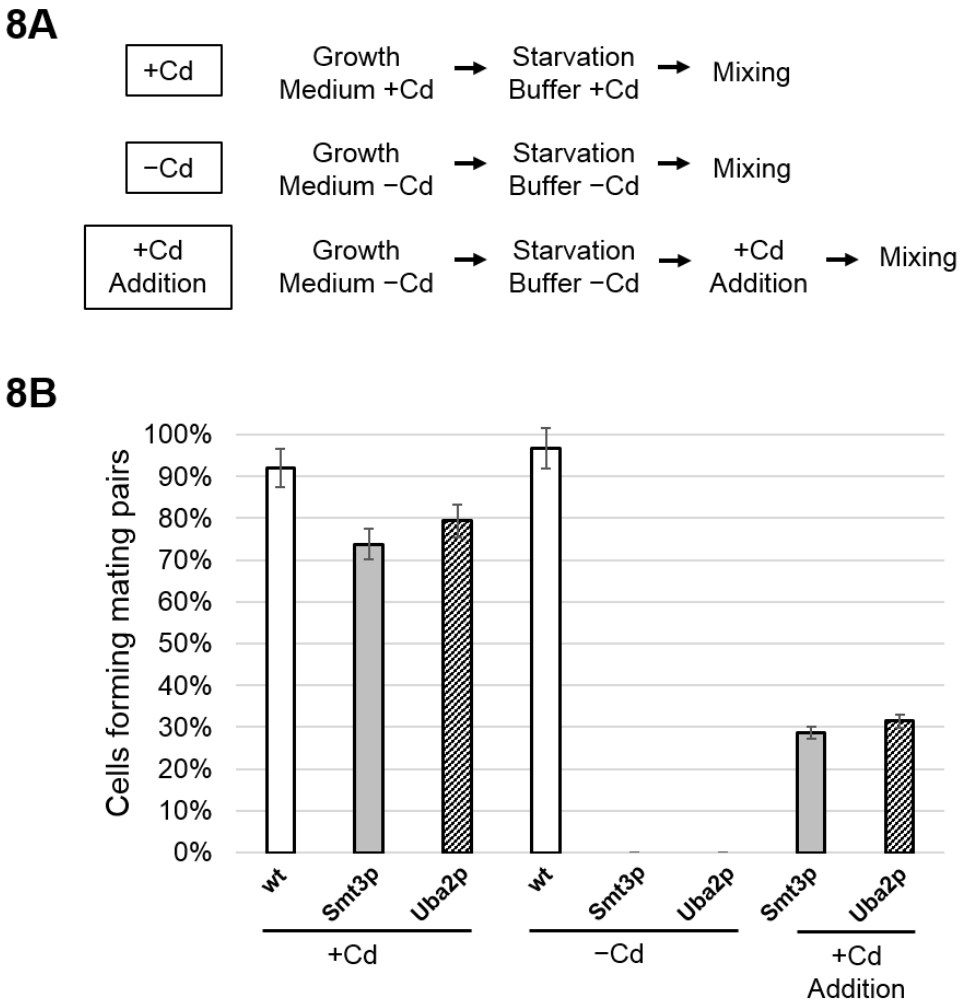

Figure 8. Depletion of Smt3p and Uba2p results in inability to form mating pairs. (A) Work flow for assaying formation of mating pairs in *SMT3* and *UBA2* conditional mutants. Strains of different mating types for each genotype were cultured with (+Cd) or without (–Cd) cadmium to early log-phase, then starved in Starvation Buffer with the corresponding plus or minus cadmium condition (see Materials and Methods for details). A portion of each –Cd culture was supplemented with 0.1 µg/mL cadmium two hours prior to mating, these cultures are referred to as "+Cd Addition". Cells were then mixed to initiate mating. (B) At 8 hours post-mixing, cells were fixed and scored using a hemocytometer to calculate mating efficiency. *SMT3* and *UBA2* conditional mutants cultured in cadmium-deficient conditions (–Cd) were unable to form mating pairs and did not progress through conjugation. Cells from +Cd cultures or –Cd cultures that were supplemented with cadmium two hours prior to mixing (+Cd Addition) formed mating pairs. Solid white bars represent wild-type (wt) cells, grey bars represent *SMT3* conditional mutant (Smt3p) and hatched bars represent *UBA2* conditional mutant (Uba2p). Error bars (standard error) are shown.

Smt3p and Uba2p deficient strains are sensitive to DNA damaging agents

Generation of conditionally mutant strains further allowed us to examine possible roles of this modification. Studies in budding yeast have shown increased DNA damage sensitivity correlated with reduced SUMOylation (121–123). Multiple studies in mammals show Smt3p modification of DNA repair proteins enhances the repair response (75, 76). To test whether *Tetrahymena* relies on SUMOylation-mediated DNA repair, we treated Smt3p and Uba2p conditional cell lines with the DNA-damaging agents: methyl-methanesulfonate (MMS) that causes cell cycle arrest and cisplatin that cross-links DNA causing cell cycle arrest and induces apoptosis in yeast and mammalian cells (124–126). Wild-type, Smt3p conditional lines and Uba2p conditional lines were cultured in the presence or absence of cadmium. Cells were treated with either MMS or cisplatin, and then placed in drops with cadmium (0.1 μg/mL) for 24 hours to evaluate cell survival. In the absence of cadmium, Smt3p and Uba2p conditional mutants showed lower survival against MMS and cisplatin compared with the same cells cultured with cadmium (Figure 9). This increased survival after cadmium exposure occurs despite the negative effect of cadmium on wild-type cell survival (Figure 9). This finding suggests that SUMOylation enhances the DNA damage repair response in *Tetrahymena* as it does in other species. The role of increased SUMOylation in conferring resilience to DNA damage in conditional mutants is consistent with a role for SUMOylation during the programmed DNA repair events of MAC development.

9

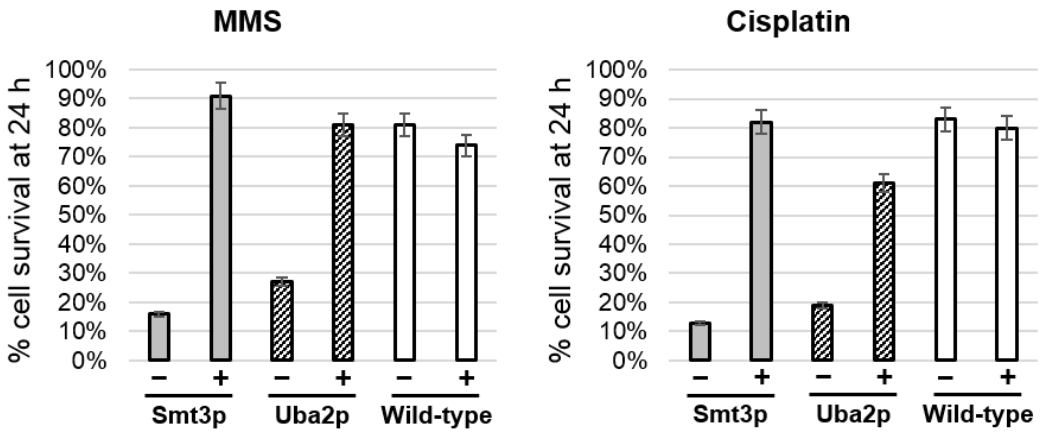

Figure 9. Smt3p and Uba2p depleted cell lines are sensitive to DNA damaging agents. *SMT3* and *UBA2* conditional mutants growing in the absence or presence of cadmium (1 µg/mL) were treated with DNA damaging agents MMS (4.5 mM) and cisplatin (2 mM). For each condition 216 single cells were isolated into drops with medium supplemented with 0.1 µg/mL cadmium and incubated at 30 °C for 48 hours. Drops containing ≥500 cells were scored as viable and drop containing ≤10 cells were counted as unviable. Viability is expressed as percentage of viable drops out of total number of drops. B2086 was used as the wild-type control. Error bars (standard error) are shown.

Discussion

Developmentally regulated SUMOylation during conjugation in Tetrahymena

Global increases in SUMOylation are known to result from exposure to environmental conditions or specific reagents, for example hydroxyurea, heat shock or ethanol (47, 59, 76). In contrast, few examples of large developmentally regulated increases in SUMOylation have been reported. Our study reveals that the sexual cycle of *Tetrahymena* is one such example. Immunoblot analysis of whole cell extracts revealed differential modification of substrates by Smt3p between vegetative and mating *Tetrahymena* cultures (Figure 4). Smt3p increased as conjugation progressed with the highest Smt3p signal observed during MAC differentiation stage. This increase is consistent with elevated *SMT3* and *UBA2* transcript levels in conjugating cells compared with vegetative or starved cells (111). Our immunofluorescence studies also showed that Smt3p is predominantly a nuclear protein during the MAC development stage (Figure 5). Together, the immunoblot and immunofluorescence data suggest a requirement for protein SUMOylation inside the developing MAC. During this period extensive genome remodeling occurs including heterochromatin formation, transcriptional regulation, DNA replication and DNA repair (96, 108), processes that have known connections to SUMOylation in other species. In addition, our previous study of *UBA2* and *SUMO* genes in *Paramecium* revealed elevated transcript levels for both during conjugation. RNAi generated knockdowns arrested conjugating cells and inhibited programmed DNA elimination (111). While the nuclear events in *Tetrahymena* and

Paramecium are dramatic, other reports of global changes in SUMOylation are known. During mouse spermatogenesis two distinct expression profiles of SUMO isoforms were detected, one during meiosis and the other in post-meiotic spermatids suggesting an important role for SUMOylation in spermiogenesis (86). A recent study in human keratinocyte differentiation also highlights SUMOylation as a regulator of cell differentiation. In these cells, SUMO expression was upregulated by calcium signaling at both the RNA and protein level, while degradation of SUMO activating and conjugating enzymes resulted in abnormal differentiation of these cells demonstrating key roles for SUMOylation in the keratinocyte differentiation process (127).

The SUMO pathway is conserved in Tetrahymena

Despite the dramatic changes in SUMOylation in *Tetrahymena* we find that most features of the pathway are conserved with other eukaryotes. As expected based on other species, the localization of Smt3p and Uba2p is predominantly nuclear (reviewed in (1)). In vegetative *Tetrahymena* cells both Smt3p and Uba2p localized to the somatic MAC (Figure 3). Interestingly we did not observe a GFP-Uba2p or GFP-Smt3p signal in the mic nor did we detect the loss of mics in our conditional cell lines. We examined more than 50 cells during various stages of mitosis and could not detect a GFP-Smt3p signal from the micronucleus (data not shown). We cannot claim that Smt3p is completely absent from the mic, but our data support a much lower concentration in mics than in MACs during vegetative growth.

As true for other eukaryotes, *UBA2* and *SMT3* are essential genes. We generated germ-line deletions of *SMT3* and *UBA2* genes using homologous recombination. The deletions were carried in the mics of heterokaryon cell lines that express wild-type *SMT3* and *UBA2* from the somatic MAC. The progeny of these heterokaryon cells have complete deletions of *SMT3* or *UBA2* and die following conjugation (Table 2). This lethal phenotype is consistent with the deletion of *SUMO* and *UBA2* genes in yeast (32, 33) and mammalian model systems (81). Interestingly, our earlier work with the ciliate *Paramecium* showed that RNAi knockdowns of *SMT3* and *UBA2* had no effect on vegetative growth (111). In retrospect this lack of vegetative phenotype is likely due to incomplete knockdowns providing a low level of *SMT3* and *UBA2* expression. As observed for other eukaryotes, *Tetrahymena* Smt3p and Uba2p depleted cell lines were hypersensitive to DNA damaging agents. When we treated *Tetrahymena* with MMS or cisplatin, cells depleted for Smt3p and Uba2p showed increased sensitivity to both DNA-damaging agents (Figure 9). Several proteins involved in DNA repair are SUMO substrates or proteins that interact with components of the SUMO pathway (75, 76, 128). Examples include DNA repair proteins such as PCNA (122), 53BP1 and MDC1 (75) and XRCC4 (128). 53BP1 and MDC1 localize along with SUMO proteins to sites of double strand breaks that occur as a result of DNA damage (75). XRCC4, an important protein in the mammalian NHEJ pathway, depends upon transient SUMOylation for localization to the nucleus (128). The NHEJ pathway is believed to be the key mechanism for repair of dsDNA breaks generated during programmed DNA rearrangements that occur naturally

during conjugation in ciliates (129). In *Paramecium tetraurelia* the RNAi induced silencing of ligase IV (a partner of XRCC4 in NHEJ) during conjugation results in the persistence of free broken ends during genome reorganization (130). The increasing links between NHEJ pathway proteins and genome reorganization during ciliate conjugation suggests potential targets for SUMOylation. The increased sensitivity of conditional Smt3p mutants to DNA-damaging agents along with the accumulation of Smt3p and Uba2p in the MAC anlagen is consistent with a role for SUMOylation during MAC differentiation in ciliates.

SUMOylation pathway is required for cell pairing

When we reduced Uba2p expression levels in conditional mutant strains by withdrawing cadmium, the cells were unable to form mating pairs. The same conditional strains growing in the presence of cadmium proceeded through conjugation normally as did wild-type cells. Although this result was unexpected, there is precedence for SUMOylation dependent effects on mating. In the budding yeast, SUMO-modification of transcription factor Ste12 is stimulated by mating pheromone thus increasing its half-life and committing the cell to the mating differentiation program (131). In addition, turnover of the yeast mating type factor, α1 protein, is dependent on SUMO-targeted ubiquitin ligases SLX5 and SLX8 which is consistent with the involvement of SUMO-mediated pathways (132). In our system, the inability of Smt3p and Uba2p deficient *Tetrahymena* cells to form pairs could result from altered gene transcription, modified signaling pathways or direct

SUMOylation of the mating type protein. The experiments presented here do not allow us to distinguish between direct and indirect SUMO-mediated effects but the recent identification of the mating type protein from *Tetrahymena* provides an opportunity to test some of these possibilities in the future (133). The dramatic upregulation of SUMOylation in *Tetrahymena* coupled with the defined events of genome alteration (heterochromatin formation, DNA elimination, DNA replication) make this a rich system for analysis of SUMOylation function and dynamics. Although the cell-pairing defect prevented the use of these lines to analyze defects at later stages of conjugation, we expect strong phenotypes during macronuclear development. Efforts are currently focused on developing alternative approaches such as inducible RNAi knockdowns of SUMO pathway genes and identifying SUMOylated proteins specific to conjugation for further analysis.

CHAPTER 3 – PROTEOMIC STRATEGY TO IDENTIFY SMT3P SUBSTRATES IN *TETRAHYMENA*

Early in the study of SUMOylation, most substrates were identified by chance when immunoblot analyses revealed slower migrating bands that were unexpected. Prior to 2004, the list of known Smt3p substrates had been compiled mostly by chance discovery of SUMO modification on proteins of interest and direct candidate analysis which most commonly involves epitope tagging of proteins purported to be SUMO-modified and use of immunoblotting to confirm the SUMOylation state of these proteins. Proteomic-based methods are increasingly useful in the identification and validation of proteins modified by UBLs such as SUMO because they offer high throughput and are far less labor-intensive than testing the SUMOylation status of individual proteins. The most common approach involves expression of tagged SUMO proteins followed either by single or double affinity purification (reviewed in (92)). The use of immunoprecipitation (as one of the affinity purification steps) is becoming very common because it utilizes specific interactions between the antibody-coupled matrix and epitope-tagged SUMO protein (134). Following the affinity purification steps, protein samples are resolved by SDS-PAGE and then digested with a protease to yield short peptide fragments. Peptides are further separated using by High Performance Liquid Chromatography

(HPLC) and eluted peptides are injected into the mass spectrometer (MS) instrument. Inside the instrument, the peptides are ionized and are directed to the mass analyzer, which separates them on the basis of their mass-to-charge ratio (m/z). The most commonly used method in the detection of UBLs is tandem MS (also indicated as MS/MS) where primary intact peptides fragment into intense ions that are selected for a second round of ionization, collision and detection. This generates a series of peptidic fragments that have characteristic m/z values that are used to construct the peptide sequence of the parent ion. In the final stage MS data is analyzed by matching the experimentally observed m/z values of each peptide with the theoretical ones contained in a database. The *in silico* digest of the peptides is done by peptide-search engines such as Mascot, Sequest and X!Tandem which will look for shifts in the mass values due to addition of PTMs on protein substrates (reviewed in (93)). There are issues with proteomics-based approach in the identification of the SUMO proteome namely, low-stringency purifications, which means fewer actual protein substrates than desired are immunopurified, and ineffective protease inhibition which results in removal of the Smt3p moiety from the protein substrates resulting in loss of the tag as well as general degradation of proteins. As a result, methods that address these two primary concerns prove very useful in the enrichment of SUMOylated substrates from a complex biological mixture. Therefore, we used a denaturing cell lysis to prevent deSUMOylation of substrates followed by two-step affinity chromatography coupled with HPLC-MS/MS analysis for identification of SUMOylated substrates. Our data revealed

SUMO-modified proteins controlling chromatin structure, gene expression, metabolism, and protein translation.

Materials and Methods

Generation of Tetrahymena expressing His6-3FLAG epitope-tagged SMT3

The double epitope construct consisted of a hexahistidine (His6) tag followed by a triple FLAG (3FLAG; FLAG tag sequence is DYDDDK) tag preceding *SMT3* gene. The tagged construct was inserted into the unique BsiWI and ApaI sites of pBSYFP2_gtw plasmid (Douglas Chalker, Washington University, St Louis) under the control of the *MTT1* promoter along with a selectable marker in the form of a mutant *rpl29* gene that confers cycloheximide resistance via homologous recombination. The tagged Smt3p construct was introduced into wild-type B2086 and CU428 strains as described (22). Cells successfully transformed with the *His6-3FLAG-SMT3* construct were selected with cycloheximide (12.5 μg/mL). Individual cells were isolated in drop medium to establish colonial cell lines that were subsequently induced with 0.1 μg/mL $CdCl_2$ and confirmed by Western Blot analysis for expression of the His6-3FLAG-Smt3p construct. For control experiments, wild-type untagged strains were used.

Purification of Smt3p Conjugates

Tetrahymena strains containing *His6-3FLAG-SMT3* were inoculated at 1000 cells/mL and grown in 500 mL of SPP medium to early log-phase (2×10^5 cells/mL) at 30 °C in a platform shaker (110 rpm). Cells were harvested at room temperature in 250 mL spin bottles (Corning, Corning, NY) by centrifugation at 2,250 rpm in a J2-MC High-Performance centrifuge (Beckman Coulter, Indianapolis, IN). Cells were

then washed once with 100 mL of Starvation Buffer (10 mM Tris-Cl, pH 7.5) and then pelleted using the same conditions as the initial harvest. The pellet was resuspended in 10 mL of lysis buffer (8 M Urea, 100 mM Na_2HPO_4-NaH_2PO_4, 10 mM Tris-Cl, 5 mM imidazole, pH 8.0) containing 50 mM N-ethylmaleimide (NEM) and protease inhibitors (final concentrations of 10 μg/mL chymostatin, 5 μg/mL leupeptin, 12.5 μg/mL antipain, and 10 μg/mL E-64). The resulting whole cell extract (WCE) was immediately placed on ice and further lysed by 15X pulses of sonication (50% power, output setting 7) using a Misonix XL-2015 sonicator (Qsonica, Newtown, CT). The WCE was clarified by centrifugation at 4 °C for 1 hour at 25,000 rpm in a SW41 rotor (Beckman Coulter, Indianapolis, IN) and then further clarified by passage through Miracloth (EMD Millipore, Billerica, MA) placed inside a 10 cc syringe. The filtered lysate (10 mL) was incubated with a 250 μL slurry of Ni-NTA resin (QIAgen, Valencia, CA) for 24 hours in a rotator at 4 °C. The resin was washed for 10 minutes with 10 mL each of Wash Buffer A (8 M Urea, 100 mM Na_2HPO_4-NaH_2PO_4, 10 mM Tris-Cl, pH 8.0), Wash Buffer B (8 M Urea, 100 mM Na_2HPO_4-NaH_2PO_4, 10 mM Tris-Cl, 0.2% Triton X-100, pH 8.0) and Wash Buffer C (8 M Urea, 100 mM Na_2HPO_4-NaH_2PO_4, 10 mM Tris, 0.1% Triton X-100, pH 8.0). Wash buffers were removed from columns by gravity flow (bottom stoppers opened). Bound proteins were eluted 3X with 1.5 bed volumes of Wash Buffer A supplemented with 250 mM imidazole resulting in a total volume of ~600 μL. The elution fraction was then diluted to 10 mL with cold TBS (50 mM Tris, 150 mM NaCl, pH 7.5). The diluted sample was then incubated with 100 μL of anti-DYDDDK

agarose (GenScript, Piscataway, NJ) at 4 °C for 24 hours. Beads were then pelleted by centrifuging at 1,000 rpm in a J2-MC High-Performance centrifuge (Beckman Coulter, Indianapolis, IN) for 1 minute. Beads were transferred to 1.7 mL microfuge tubes and then washed with 1 mL of TBS by rotating at 4 °C for 10 minutes. Wash steps were repeated 3X times. Beads were pelleted after each wash by spinning at 1000×*g* in a mini centirfuge. TBS Proteins were eluted twice with 250 µL of alkaline triethylamine (100 mM, pH 12.0). The control protein preparation for this experiment was prepared in the exact same manner, except untagged wild-type *Tetrahymena* strains were utilized.

Western Blotting and Coomassie Staining Analyses

To evaluate expression of the tagged SUMO lysates from the His6-3FLAG-Smt3p and control wild-type strains were prepared for SDS-PAGE by addition of 4X SDS loading buffer (63 mM Tris, pH 6.8, 2% glycerol, 2.3% SDS, and 0.05% bromophenol blue) supplemented with 100 mM dithiothreitol (DTT) to a final concentration of 1X. Protein samples were incubated for 5 min in a 95 °C heating block before loading on 8% Bis-Tris acrylamide gels (27.5:1 bis-acrylamide ratio) using Tris-Glycine SDS running buffer. SDS-PAGE gels were electrophoresed at 100 V for 2 hours. For Western Blotting, proteins were transferred to PVDF membrane under wet conditions (Pall Corporation, Port Washington, NY) at 147 V for 1.5 hours. Western blot analysis was then performed with initial probing with anti-FLAG primary antibodies (Sigma, St. Louis, MO) prepared as a 1:10,000 dilution in

5% skim milk. Primary antibody incubation was performed overnight on a platform shaker at 4 °C. Following this, membranes were probed with horseradish peroxidase-conjugated secondary antibodies (Jackson Immunoresearch, West Grove, PA), also prepared as a 1:10,000 dilution in 5% skim milk, for 1 hour at room temperature. The membranes were incubated with Luminata Crescendo detection reagent (EMD Millipore, Billerica, MA) for 2 minutes at room temperature. Results were visualized with the ChemiDoc Imager (Bio-Rad, Hercules, CA) using the default program for western blotting. Coomassie staining of Bis-Tris acrylamide gels was performed by first fixing the gel in fixing solution (10% glacial acetic acid and 25% isopropanol) for 3 hours and then staining with Coomassie R250 staining solution (0.01% Coomassie R250 and 10% acetic acid) overnight at room temperature. Destaining (10% glacial acetic acid) was done for 3-6 hours with gentle rocking.

In solution Trypsin Digestion and Protein Fractionation

Eluates from the second round of purification using FLAG resin were pooled together and digested with trypsin (1:100 enzyme:protein ratio; Piercenet, Rockford, IL) at 37 °C for 24 hours. The tryptic digestion was quenched by addition of 0.1% trifluoroacetic acid (TFA). Peptide preparations were individually loaded onto C18 resin (Thermo Scientific, Waltham, MA) and separated using a 35-min gradient from 2.5 to 97.4% acetonitrile.

Mass Spectrometry Acquisition and Database Search

Eluted peptides were then directly entered into an ESI QuadTOF mass spectrometer (TripleTOF 5600; AB Sciex, Concord, Canada). The instrument was coupled with an Eksigent Nano LC-2DPlus with nanoFlex cHiPLC electrospray ionization (ESI) system (Eksigent, Dublin, CA). The sample acquisitions were performed using a "trap and elute" configuration on the nanoFlex system. The trap column (200 m 0.5 mm) and the analytical column (75 m 15 cm) were packed with 3 m ChromXP C18 medium. Samples were loaded at a flow rate of 2 nL/min for 10 minutes and eluted from the analytical column at a flow rate of 300 nL/min in a linear gradient of 5% solvent B (95% acetonitrile, 0.1% formic acid, v/v) to 35% solvent B in 90 min. The column was regenerated by washing with 80% solvent B for 10 minutes and re-equilibrated with 5% solvent B for 10 minutes. For standard data-dependent analysis experiments, the mass spectrometer alternated between acquiring a full 250-ms scan (TOF-MS) and subsequent MS/MS spectra of the 50 most abundant precursor ions with a 50-ms scan. The selection criteria for parent ions included intensity, where ions had to be greater than 100 counts/s with a charge state between +2 to +4 and were not present on the dynamic exclusion list. Once an ion had been fragmented by MS/MS, its mass and isotopes were excluded for a period of 6 s. Ions were isolated using a quadrupole resolution of 0.7 Da and fragmented in the collision cell using collision energy ramped from 15 to 45 eV within the 50-ms accumulation time. In the instances where there were less than 20

parent ions that met the selection criteria, those ions that did were subjected to longer accumulation times to maintain a constant total cycle time.

Data were analyzed using the *Tetrahymena* Genome Database (TGD) <http://ciliate.org/> in conjunction with the Mascot v2.5 search algorithm (Matrix Science, Boston, MA). MS searches allowed for a search tolerance of 10 ppm for the peptide mass tolerance and 0.2 Da for the MS/MS tolerance. The enzyme selected for protease digestion was trypsin with 2 missed cleavages specified. The charge of the peptides to search was set to +2, +3 and +4. Initial peptide filtering was done at an FDR of 1% and proteins were required to have 2 or more peptides identified at the 1% FDR level. The MS spectra were searched against a reverse database to estimate the false discovery rate. Figures and tables of the identified proteins were made using information from both the TGD and UniProt <http://www.uniprot.org/>. Finally, the sequences of identified proteins were analyzed to locate potential SUMOylation consensus motifs using SUMOsp2.0 <http://sumosp.biocuckoo.org/>.

Results

Affinity Purification of Smt3p Conjugates from Tetrahymena

The main challenge in identifying Smt3p substrates is the low abundance of protein substrates relative to the total protein in the cell. In addition, SUMOylation is often a short-lived and transient modification which means that few proteins in the cells are likely to be SUMOylated. The use of two rounds of purification serve to enrich for SUMOylated substrates for identification in the mass spectrometer. To determine the catalog of SUMO proteins (the SUMOome) in vegetative cells, we used a two-step affinity purification followed by HPLC-coupled MS protein identification. The dual tag construct consisted of the N terminal 6His tag followed by a 3FLAG tag. We used the *MTT1* promoter for cadmium-based induction of the tagged Smt3p (Figure 10A). The construct was transfected in *Tetrahymena* strains at the *rpl29* locus that confers resistance against the drug cycloheximide. It should be noted that these cells also express wild-type Smt3p from the endogenous locus. As shown in Figure 10C expression of the tagged Smt3p was cadmium dependent. A low concentration of cadmium (0.1 μg/mL $CdCl_2$) was used to induce the expression of the dual-tagged 6His-3FLAG-Smt3p protein, to avoid inundating the cell with excess Smt3p protein. The tagged Smt3p was detected in its non-conjugated (free) form at approximately 14 kDa but the majority of signal was located in higher mass regions of the gel as expected for SUMO conjugated to target proteins.

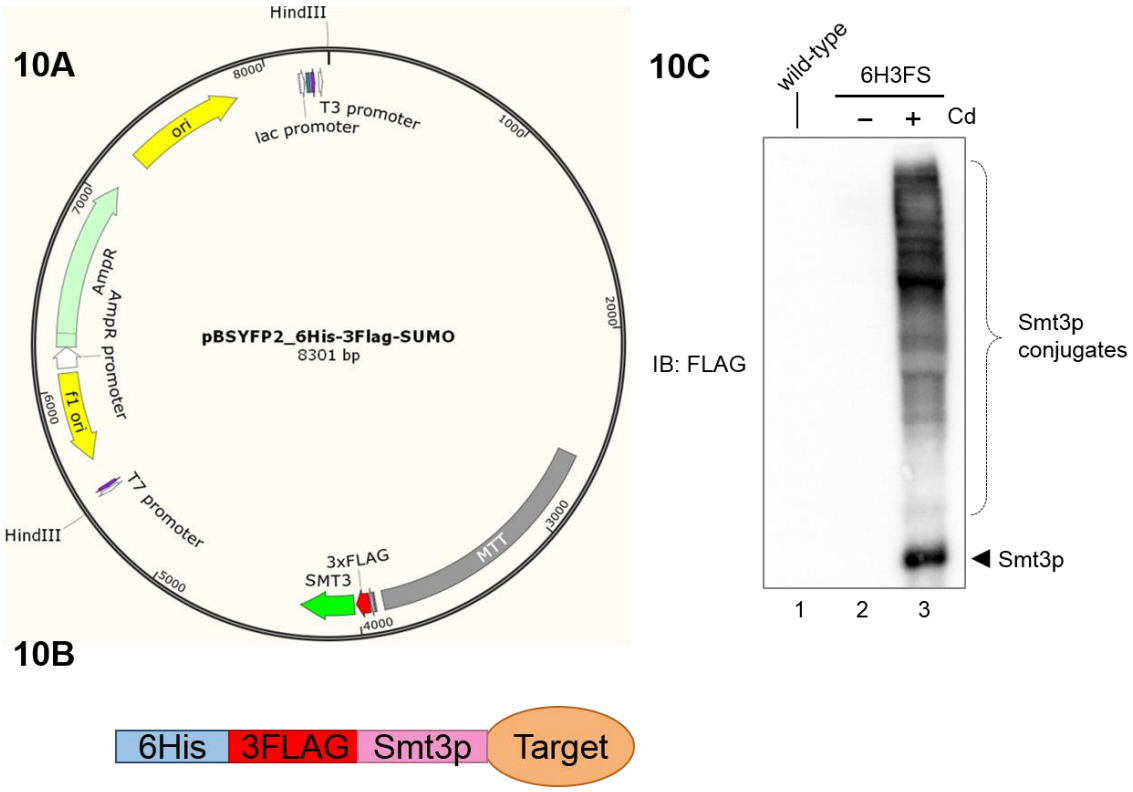

Figure 10. The psYFP2_His6-3FLAG-SUMO vector and its expression product. (A) A map of the plasmid encoding His6-3FLAG-Smt3p construct that was used to transfect wild-type *Tetrahymena* strains. (B) Schematic representation of the dual epitope tag containing the N-terminal His6 epitope tag and the second 3FLAG epitope tag. Target proteins are conjugated at the C-terminus of Smt3p. For immunoblotting analyses, antibodies directed against the 3FLAG epitope tag were used. (C) Immunoblotting of the whole cell extract (WCE) prepared from His6-3FLAG-Smt3p expressing strain and wild-type untagged strain. Expression of His6-3FLAG-Smt3p is driven by the MTTp promoter. Growth medium was supplemented with 0.1 µg/mL $CdCl_2$ to induce the expression of the tagged form of Smt3p. Free Smt3p is indicated with an arrow head at ~14 kDa. The high molecular weight region containing numerous proteins that are likely to be Smt3p substrates is also indicated with a bracket. Low amounts of cadmium were used to minimize overexpression which contributes to non-typical conjugation of protein substrates by Smt3p.

In the first round of affinity purification, the His6 tagged was bound to Ni-NTA agarose under strongly denaturing conditions (containing 8 M urea) in order to

solubilize the majority of cellular proteins by denaturation as well as to ensure that proteins identified in the proteomic screen are directly conjugated to Smt3p. The lysis buffer also contained NEM and protease inhibitors to inhibit the activity of SUMO proteases from removing Smt3p from substrates. A single round of purification is not recommended when purifying UBL substrates as it has been shown to result in a large number of false positives (109). This is especially the case with Ni-NTA purification as many proteins bind non-specifically to this resin in yeast (45, 109, 135). In order increase the purity of the sample, a second purification step was performed with anti-FLAG resin. Following incubation with anti-FLAG resin, proteins were eluted under highly alkaline conditions (100 mM alkaline triethylamine).

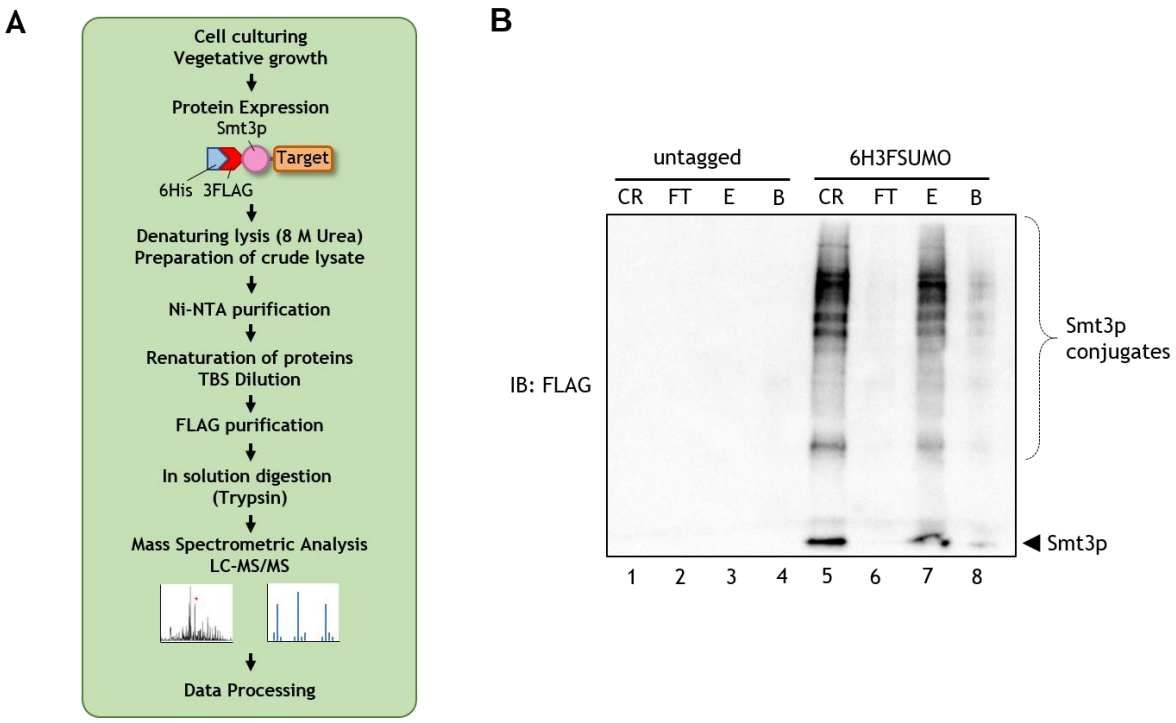

Figure 11. Overview of the double affinity purification and identification of His6-3FLAG-Smt3p tagged proteins. (A) The two-step affinity purification workflow is described in more detail in the Materials and Methods section of this chapter. The actual procedure of the double-affinity purification includes additional steps such as culturing and lysis of the cells, denaturation and renaturation of the proteins, and in solution digestion with trypsin and database searches. (B) Fractions from Ni-NTA purification were collected after each step during purification as the crude cell extract (CR), flow through (FL), elution (E), and beads (B). Samples were loaded proportionally and resolved by SDS-PAGE (10% gel). Fractions from untagged wild-type strain are on the left half (lanes 1-4) while proteins from the tagged Smt3p strain are on the right (lanes 5-8). Smt3p conjugates and free Smt3p protein are indicated.

Identification of SUMOylated Substrates by LC-MS/MS

Eluted samples from the two-step purification were digested in solution with trypsin. The resulting peptides were extracted and loaded onto a reverse-phase column and eluted with increasing concentrations of acetonitrile (ACN) directly into the mass spectrometer. We chose a QuadTOF mass spectrometer to identify the

protein substrates because this instrument has high sensitivity and accuracy. The mass spectrometer was configured to acquire both accurate mass (MS) and sequence (MS/MS) information for the eluting peptides. We used Mascot algorithm to identify the peptides against a conceptual translation of *Tetrahymena* gene sequences from Tetrahymena Genome Database (TGD) and Uniprot. Proteins that were included in our list of SUMO substrates had to meet two criteria: 1) they must be unique to the Smt3p immunoaffinity purification (not in the untagged cell purification) and 2) they must be identified by 2 or more peptides (Table 3 and Table 7). Proteins that were identified in the control wild-type untagged sample were eliminated from the tagged Smt3p substrate dataset even if they met the criteria stated above for MS search. Proteins in the untagged cell sample could be highly abundant proteins that were carried over during the two purifications. As predicted from the above-mentioned immunoblot, little overlap was observed between the datasets of the tagged Smt3p and untagged purifications.

Analysis and Categorization of SUMOylated Substrates

A total of 153 proteins were identified from five independent runs in the His6-3FLAG-Smt3p purification (full list available in Appendix Table 7). We identified numerous novel proteins as Smt3p substrates as well as others that have been previously identified (Table 3). Many of these proteins are orthologous to well characterized SUMO substrates found in other organisms. For example, Heterochromatin Protein 1 (HP1) is a SUMO substrate (136, 137) and the ribosomal

protein Rps17 has been purified from a SUMO proteomic screen in *Drosophila* (48) (Table 3).

Table 3. List of selected Smt3p substrates identified in vegetative *Tetrahymena*.

Identified Proteins	Molecular Mass (kDa)	TGD ID	References
SUMOylation Machinery			
Smt3p*	10.2	TTHERM_00410130	(138, 139)
Ubc9p*	24.0	TTHERM_00522720	(140)
Metabolic Enzymes			
Gpm1*	31.6	TTHERM_00641240	(141, 142)
Cit1*	63.6	TTHERM_00529790	(143)
Pgi1*	63.6	TTHERM_01108700	(45)
Tcb2	24.7	TTHERM_00068170	
Ribosomal Proteins			
Rpl20	29.2	TTHERM_00476670	
Rpl43	11.4	TTHERM_00075670	
Rps11	18.1	TTHERM_01109770	
Rps17*	14.9	TTHERM_00762890	(48)
Rack1*	38.5	TTHERM_01113100	(144)
Chromatin-Associated			
HTA3*	15.3	TTHERM_00143660	(145)
TGP1*	83.4	TTHERM_00499420	(146)
HP1*	44.2	TTHERM_00245410	(136, 137)
Cytoskeleton Structure			
DFB1	48.1	TTHERM_00128280	
TTNA	98.9	TTHERM_00006320	
MYO2	210.9	TTHERM_01035740	
Other			
ATP2	53.4	TTHERM_00585260	
Hsp60*	61.6	TTHERM_00196370	(147)
PRE10*	27.6	TTHERM_00487110	(109)
Liver F-Antigen	46.0	TTHERM_00678260	

*Previously identified as SUMO substrates in other species.

We compared our tagged Smt3p dataset with separate studies in *S. cerevisiae* (45), *D. melanogaster* (48) and *A. thaliana* (148) that utilized a similar two-step

affinity purification (Figure 12). There was only one protein in common with the *Arabidopsis* study and this is likely due to the fact that affinity purification in that study yielded only 20 proteins. A total of 11 proteins were common between *Tetrahymena* and *Drosophila* and 16 proteins were common in the purifications from *Tetrahymena* and yeast. There are 3 proteins that are common all three purifications. The presence of *Tetrahymena* proteins in purifications from *Drosophila* and yeast adds confidence to these proteins as authentic Smt3p substrates.

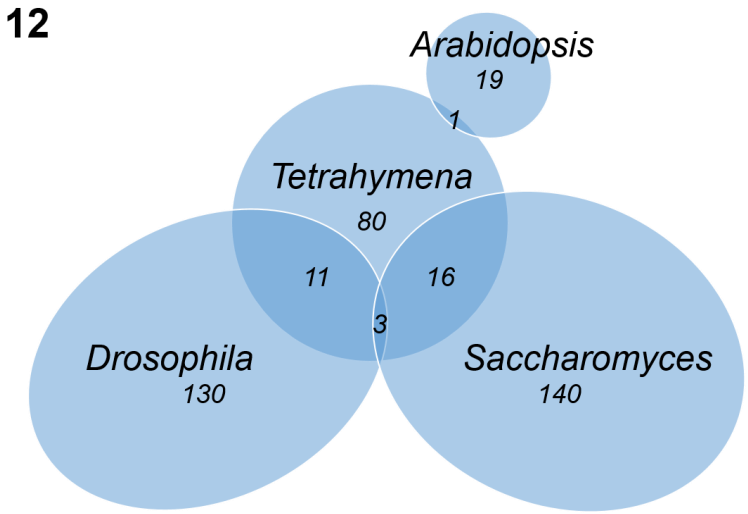

Figure 12. Comparison of SUMO substrates identified using 2-step purification in various model systems. Tagged Smt3p dataset from two-step purification in *Tetrahymena* were compared to the same from purifications from *Drosophila* and *S. cerevisiae*. Three proteins are common between all three organisms making them highly likely to be true Smt3p substrates.

In a separate study, tagged Smt3p conjugates were isolated under native conditions using the same two-step affinity purification scheme, and 141 protein substrates were identified (Qianyi Yang and James Forney unpublished). There is

some overlap between this dataset and the one described above where Smt3p conjugates were isolated under denaturing conditions (Figure 12). These protein substrates that were purified from both lysis conditions are considered high confidence Smt3p targets (Table 8).

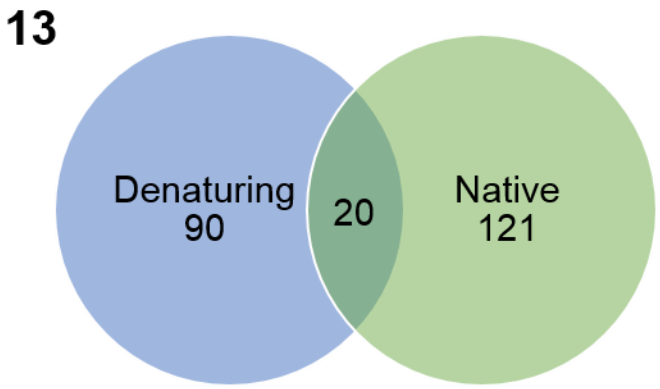

Figure 13. Overlap between Smt3p substrates purified under denaturing and native conditions. All proteins reported to be SUMOylated in our proteomic screen using denaturing conditions were manually extracted and compared to those found by MS using native conditions. Proteins considered in the denaturing lysis purification analysis are included in the Table 7 and proteins common to both lysis conditions are presented in Table 8. 20 proteins are purified from both lysis conditions indicating that these proteins are high confidence protein substrates of Smt3p.

In an attempt to gain insights into the functions of potential Smt3p substrates identified in our proteomic screen for vegetative *Tetrahymena*, we used the *Tetrahymena* Genome Database (TGD) and the Uniprot Knowledgebase to categorize these proteins based on their subcellular localization (Figure 14A) and biological functions (Figure 14B).

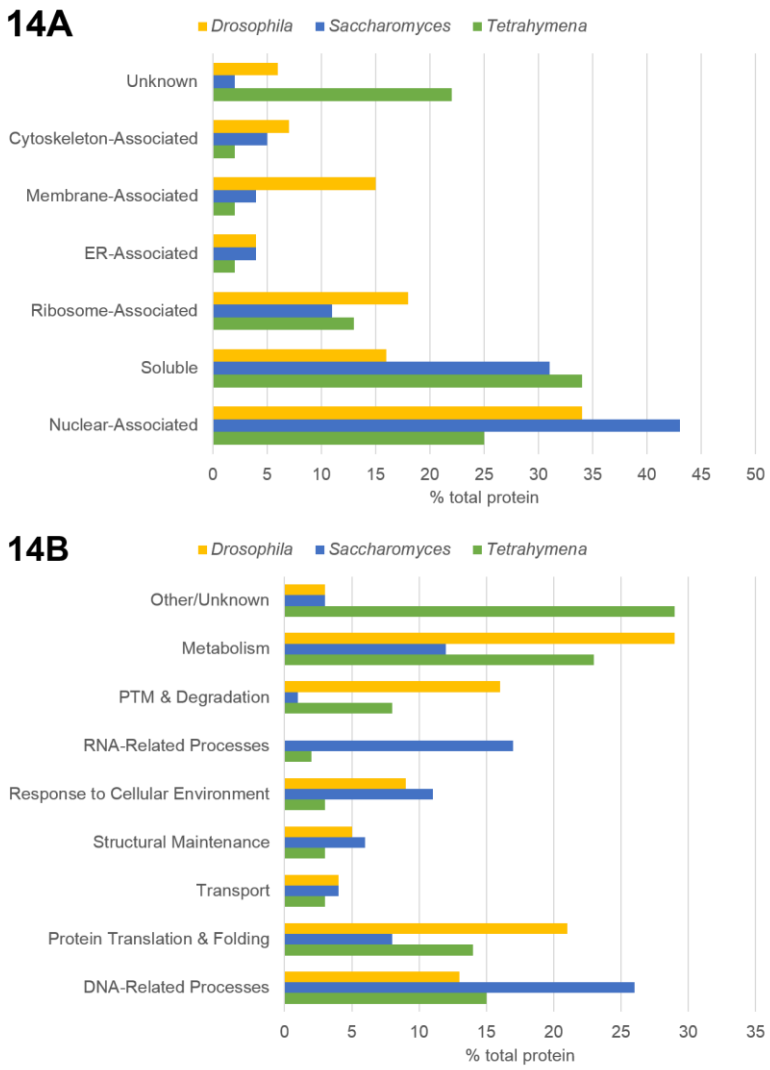

Figure 14. Graphical representation of both the cellular location and biological function of Smt3p candidate substrates. (A) We used the *Tetrahymena* Genome Database (TGD) and the Uniprot Knowledgebase to categorize the identified substrates based on cellular location. As expected, many of these proteins are nuclear proteins. However, a surprisingly high number of these proteins that are present in the cytosol and/or localized to other non-nuclear organelles. (B) These same database sources were also utilized to graphically represent the diverse functional categories (Gene Ontology terms) to which these identified SUMOylated substrates belong. The largest such category was found to be comprised of proteins involved in metabolism which suggests that SUMOylation is not highly centralized in the nucleus. However, SUMOylation clearly plays a role in other biological phenomena, including DNA-related processes, protein translation and folding, crosstalk with other PTMs etc. As in the case of Figure 14A, there are many proteins with unknown function and as more information is made available in the TGD, the fraction of proteins categorized as such is likely to decrease.

From this analysis, a few conclusions can be made. Firstly, as expected the largest localization category is the nucleus (25% of proteins identified in our study are predicted to reside in the nucleus). Both yeast and fruit fly have a higher percentage of nuclear substrates, 43% and 33% respectively. Second, we also observed a large percentage of proteins that are predicted to be cytosolic, or associated with other organelles which would suggest that SUMOylation is a less nuclear centralized process than previously thought. For example, 22% of the proteins were localized in soluble fractions while 13% and 9% are found in the ribosome and mitochondrion respectively. It is plausible that some of the identified non-nuclear proteins may translocate to the nucleus upon modification by Smt3p. A third conclusion is that protein substrates of SUMOylation belong to broad and diverse range of biological processes which is expected for this important PTM. In addition to DNA-related processes (15% in *Tetrahymena*), a substantial number of proteins in other cellular processes were also identified as targets of SUMOylation. For example, numerous proteins in *Tetrahymena* are involved in various metabolic pathways (23%) and the same is the case in *Drosophila* (28%). We also observed several proteins (14% in *Tetrahymena*) that play a role in protein translation (as part of the larger ribosome complex) similar to both yeast and *Drosophila*. Other processes include crosstalk with PTMs (9%), response to cellular environment (3%) and structural maintenance (3%). Interestingly, components of the SUMO pathway such as the E2 conjugating enzyme Ubc9p and Smt3p protein itself were observed in

our proteomic screen consistent with similar observations in other systems (Table 3) (149–151).

For novel candidate SUMOylated substrates, we searched the amino acid sequences using SUMOsp2.0 (152) for matches to a SUMO target consensus sequence (ψKXE) present in many SUMOylated proteins (reviewed in (2)) (Table 4). While SUMO conjugation of substrates does not always occur at SUMO consensus motifs in previously identified cases (reviewed in (42)), presence of these motifs suggests strong candidate lysines in the substrates that might be modified and, hence, a starting point for subsequent site-directed mutagenesis studies aimed at identifying the biological effect of SUMOylation in desired target proteins.

Table 4. Consensus Smt3p target sites.

Identified Proteins	**Smt3p Consensus Site (ψKXD/E)**
Heterochromatin protein 1	VKWE, VKQD, VKQE
CHROMOdomain protein 2	VKRE
C2H2 (zinc-finger protein)	PKIE, IKKE
Ubiquitin-60S	VKHD
Phosphoglycerate mutase 1, GPM1	KKQE
HSP70	FKVE, LKDE
HSP60	LKGE
SSA5 (DnaK protein)	FKVE, MKDE
CIT1	LKHE
ATP synthase	MKVE, LKFE
RACK1	FKVE
Sd2e domain protein	IKQE, LKQE, LKHE
Ubc9p	LKQE, LKDE
DFB1	VKNE, LKSE, LKDE, LKTE, IKNE, LKEE
TTNA	LKEE

Sequence analysis also revealed that our SUMOome contains a comparable number of ψKXE SUMO modification consensus motifs per protein to that contained in previously reported SUMOylation targets from other species (Table 5). The wild-type internal rejects (untagged strain) list contains almost three-fold fewer SUMO consensus sites per protein than the His6-3FLAG-Smt3p proteome (0.6 versus 2.2) and is similar to the fraction in the total *Tetrahymena* proteome (0.6 versus 0.8) (Table 5).

Table 5. Frequency of Smt3p consensus sites in known SUMO target proteins compared with this study. Smt3p modification consensus sites of the form ψKXE were identified using SUMOsp2.0 (152) set to high threshold.

Data Set	**Number of proteins**	**Predicted consensus sites**	**Proteins without consensus**	**Sites per protein**
Published Smt3p substrates	265	517	49	2.0
His6-3FLAG-Smt3p proteome[a]	60	132	64	2.2
Tetrahymena proteome	16,102	12,771	9,440	0.8
Wild-type internal rejects[b]	37	21	9	0.6

[a]Proteins identified in this study
[b]Proteins purified from the untagged strain

Discussion

Our proteomic screen employed double-affinity purification coupled to mass spectrometry (MS) to identify protein substrates. This approach generates improved purity and lower background than single affinity purifications. In experiments where Smt3p substrates were purified utilizing a single round of purification (45, 109, 135), many of the identified proteins were also purified from a mock untagged strain which would suggest that a single round of purification do not produce a pure sample for use in MS. Using a double-affinity purification scheme is then more suited for identifying Smt3p substrates as it yields cleaner sample and control preparations.

A second important feature of our approach that adds confidence to our SUMOome dataset is the use of a QuadTOF mass spectrometer for the collection of the MS/MS spectra. The advantage of this particular instrument is its high sensitivity (femtomoles). This is critical because only a small fraction of total proteins in the cell are SUMOylated at any given time. The high sensitivity of the LTQ mass spectrophotometer is helpful then, to identify the low amounts of SUMO-modified peptides.

We concede that even with our careful controls, there are likely to be some *Tetrahymena* proteins that may be falsely identified as Smt3p substrates. An example of this would be the heat shock proteins that we identified in only one experiment; these proteins may have been purified because of stress suffered by cells immediately prior to lysis. Other examples could be some of the ribosomal

proteins as well as proteins involved in heavy-metal metabolism which may have been identified in our proteomic screen because we used cadmium-driven *MTT1* promoter for expressing the tagged Smt3p. Despite using denaturing conditions (8 M urea) for lysis of cells, it is possible that some of the proteins identified in our proteomic screen are carried through the double-affinity purification because they form highly stable complexes with protein substrates that are attached to Smt3p. These concerns can be addressed by immunoblotting studies with antibodies against protein substrates to confirm the SUMOylation state of these proteins and/or by identifying lysine sites of Smt3p attachment (See *Future Directions*).

The more obvious targets of Smt3p in our proteomic screen are nuclear proteins that are involved in gene expression. Many of these protein substrates play a role in chromatin remodeling, transcription and mRNA processing (Table 3). This is consistent with a role of SUMOylation as a predominantly nuclear process (reviewed in (2)).

Our proteomic screen did not provide full coverage of the entire SUMOome. This is evident by the fact that we did not identify proteins such as topoisomerase II (4) and RanGAP1 (153) which are well known Smt3p substrates in yeast. Optimization of the double-affinity procedure to minimize high abundant sticky proteins may help in getting a purer sample. A possible solution is to separate the nuclei from the cell using differential centrifugation and analyze the nuclear extract using mass spectrometry. Lastly, it is possible that these well-known Smt3p substrates are low abundance proteins in *Tetrahymena* making it difficult to carry

these proteins over two purification steps. While our proteomics screen did not cover the entire SUMOome, we were able to identify several novel protein substrates of SUMOylation in *Tetrahymena* such as the ER protein, TCB2 and the deep-fiber bundle protein, DFB2. Similar studies have been performed in mammals, *Drosophila* and yeast using similar proteomic approaches (45, 46, 48) where full coverage of the SUMOome was not achieved. It may require multiple runs as well other proteomic approaches to produce a more comprehensive list of Smt3p targets.

SUMOylation in DNA-related Processes

Many protein substrates of SUMO are nuclear proteins with important roles in the regulation of transcription, maintenance of chromosome and chromatin structure and DNA repair (reviewed in (154, 155)). The effect of SUMOylation on transcription factor activity generally leads to repression as opposed to Ub-modification which tends to result in transcriptional activation (reviewed in (154)). As observed in proteomics studies in yeast (45) and mammals (156), the largest group of proteins identified in SUMO proteomic screens were involved in DNA-related biochemical processes. The same trend was observed in our proteomic screen. Chromatin proteins such as the histone H2A variant (HTA3) were identified as Smt3p substrates. Another nuclear protein (TTHERM_00245410) that contains a CHRromatin Organisation MOdifier (CHROMO) domain was also captured five times in our proteomics screen. This is a homolog of a heterochromatin protein (HP1/CBX1) in mammals that recognizes and binds histone H3 tails that are

methylated at Lys-9 (H3K9) leading to a heterochromatic state through epigenetic repression (136). It is known to interact with SUMO-1 in affinity-capture experiments (137). One proposed model is that SUMOylation of CBX1 causes the protein to localize to the nucleus. Mutants that cannot be SUMOylated are unable to localize in nuclear domains and are instead found dispersed throughout the cell (136).

As many of the SUMOylated proteins identified to date are transcription factors, coactivators, or corepressors that associate with DNA (reviewed in (89, 91, 154)), it is not surprising that our affinity purification captured several nuclear protein substrates highlighting the importance of this PTM in DNA-related processes. Transcriptional factors and other chromatin-associated proteins such as histones have been demonstrated to be modified by other PTMs such as phosphorylation, acetylation, methylation, and ubiquitylation (reviewed in (154)). As a result, analysis of SUMOylation-dependent changes in chromatin dynamics is undoubtedly made complicated by the fact that many of the enzymes that regulate chromatin structure are themselves modified by several PTMs. Identification of proteins involved in DNA-associated processes as Smt3p substrates aids in understanding the crosstalk between different PTMs.

SUMOylation directs Ribosome Maturation

The control of ribosome biogenesis is a critical process which ensures that protein synthesis is coordinated with cell growth and proliferation. Several proteomic studies aimed at identifying SUMO substrates have shown that ribosomal proteins and trans-acting factors have been identified as potential Smt3p substrates in various model systems (45, 48, 109, 157). We too, identified numerous ribosomal subunit proteins that are part of the greater ribosomal complex in our proteomic screen - RpS4, RpL15, RpL23, RpL4, RpL18, RpL22, RpL40, and RpL36 were all purified as Smt3p substrates. The presence of several ribosomal subunits would suggest that the multi-protein ribosomal complex often contains multiple SUMOylated proteins (Table 3 and Table 7) some of which possess several predicted sites of Smt3p attachment (Table 4). It is not entirely clear why so many ribosomal proteins are substrates of SUMOylation. Smt3p could be part of a nuclear surveillance system that ensures proper protein translation. Indeed, under conditions of stress that require increased protein translation for stress response, an overall increase in SUMOylation is observed which may help in ribosomal maturation, assembly and function (47, 76). This role of SUMOylation is supported by mutation studies where inhibition of SUMOylation affects ribosome export (157). The ribosomal protein RpL40, involved in regulation of protein translation, was also identified as a substrate of Smt3p. RpL40 has two components - ubiquitin and ribosomal 60S subunit. The Ub component of RpL40 is believed to facilitate the assembly of the ribosomal protein into ribosomes upon export to the nucleus (158).

In a recent study, investigators used a specific protease inhibitor as a way of inhibiting the activity of SENPs proteases in prostate cancer cells stably expressing SUMO (159). A combination of stable isotope labeling with amino acids (SILAC) quantitative proteomic technique was used to identify more than 900 putative target proteins of SUMO, one of which was RpL40 (159). Regulation of RpL40 by Smt3p modification may shed light on the mechanism underlying prostate cancer progression.

As demonstrated in Table 3, our work provides a few examples of clustering of SUMOylation in the ribosomal complex. The high degree of incidence of this phenomenon in our data strongly indicates that substrate clustering may be an important mechanism by which SUMOylation regulates large protein complexes. An alternate model of Smt3p regulation may be that it regulates cellular processes by targeting macromolecular complexes rather than individual proteins as typical of other posttranslational modifications. Such coordinated SUMOylation could function to localize an entire protein complex to a specific subcellular location or stabilize the protein complex by inhibiting modification by Ub. Clearly, additional work aimed at elucidating the functional strategy underlying this mechanism will provide important insight into how SUMOylation regulates these larger macromolecular complexes.

SUMO in Metabolic Homeostasis

The first identified targets for SUMO modification were mostly nuclear proteins (reviewed in (82)). However, more recent studies have identified a number of cytoplasmic Smt3p targets (reviewed in (160, 161)). Extra-nuclear roles of SUMOylation have been described in regulation of G protein signaling, phosphorylation crosstalk, mitochondrial fission/fusion, glucose transport as well as regulation of neuronal integrity and synaptic function (41, 162, 163). In some cases SUMOylation regulates master proteins to affect entire metabolic pathways (reviewed in (160)). Phosphoglycerate mutase 1 (GPM1) (Table 3), which participates in the glycolysis cycle, was identified as an interacting partner of Smt3p through a bimolecular fluorescence complementation assay in yeast (141). The same protein, GPM1, was also identified as a SUMO substrate in a separate proteomics screen in yeast (164). In *Drosophila*, a genome-wide screen to systematically search for modulators of citrate synthase (CS) identified Smt3p as an active regulator of the CS enzyme (143). Flies that are *SMT3* heterozygous mutants have increased CS activity, providing additional support for SUMOylation in regulating mitochondrial function. Considering that SUMOylation greatly influences protein activity, stability and protein interactions, it is not surprising that SUMOylation regulates the function of metabolic enzymes within pathways.

FUTURE DIRECTIONS

RNAi based knockdown of SUMOylation pathway genes

The role of SUMOylation in regulating diverse critical processes in the eukaryotic cells is well documented (reviewed in (2, 82, 91)). We observed that there is an increase in the levels of Smt3p during *Tetrahymena thermophila* conjugation (Figure 4). Conditional mutants of Smt3p and Uba2p, in which Smt3p and Uba2p was depleted, were unable to form mating pairs and therefore could not progress through conjugation. This non-pairing phenotype prevented us from analyzing the requirement of SUMOylation during conjugation. To circumvent this issue, we can use an RNAi-based approach to knockdown the levels of SUMOylation pathway genes. Previous studies in another ciliate *Paramecium tetraurelia* demonstrated that RNAi-based silencing of Smt3p and Uba2p inhibited the excision of IES elements during the macronuclear development stage of conjugation in this organism (111). We anticipate that RNAi-mediated downregulation of SUMOylation pathway genes in *Tetrahymena* will lead to a similar phenotype with cells arresting during MAC development stage.

Identification of SUMOylation sites on protein substrates

An alternate approach to study the requirement of SUMOylation during conjugation (or vegetative life cycle for that matter) is to identify and alter SUMOylated substrates. We intend to employ a MS-based approach to identify the catalogue of proteins that are modified differentially by Smt3p in response to the sexual life cycle and the various stages therein. Proteomics approaches which aid in identification of sites of SUMOylation are especially helpful in this case as they yield useful information for mutational analyses where sites for Smt3p attachment can be inactivated in protein substrates. To this end, we have generated a *Tetrahymena* strain expressing Smt3p V85K that can be used in conjunction with mass spectrometry to map SUMOylation sites on purified proteins. It is technically challenging to identify sites of SUMO attachment to protein substrates because common proteases such as trypsin generate long SUMO peptides attached to target lysines that add to the mass of the peptide and are subject to fragmentation during MS/MS. The resulting fragmentation patterns in the mass spectrum are very complex and cannot be easily interpreted by conventional MS algorithms. The Smt3p mutant with V85K mutation leaves a GG remnant on target lysine residues on substrates when affinity-purified substrates are subjected to protease digestion with Lys-C. In this case, the MS instrument can be configured to detect the specific mass shift because of attachment of the diglycine remnant to proteins substrates (46). Preliminary data shows that we have identified SUMOylated lysine residues on several proteins from both vegetative and mated *Tetrahymena* Smt3p V85K strains.

For protein substrates of Smt3p identified earlier that are not purified using the V85K mutant, we will have to use *in vitro* and/or *in vivo* methods to confirm their SUMOylation status.

HP1 and HTA3 as potential Smt3p substrates

A CHROMO domain-containing protein related to Heterochromatin Protein 1 (TTHERM_00245410, here referred to as HPR) was identified in our two-step purification procedure in *Tetrahymena* (Chapter 3). HP1 was also identified as a SUMO target in yeast. The hinge domain of murine HP1 has been demonstrated to be modified by SUMO (136). In this particular study, the authors observed that *de novo* SUMOylation of HP1 targets the protein to pericentric heterochromatin. Microarray expression profile shows that the levels of HP1 in *Tetrahymena* are upregulated during conjugation with the peak during MAC development (~3-fold higher than vegetative growth). We know that major regions of the genome in the developing macronucleus for heterochromatin before their elimination. My model for the role of HP1 in MAC development involves *de novo* targeting HPR to these heterochromatic regions followed by methylation of histone H3 lysine 9 and lysine 27 to stabilize the heterochromatin and allow CHROMO domains to bind the methylated lysines. Subsequent events include the formation of sub-nuclear structures that contain protein machinery required for DNA rearrangement.

This model makes a clear prediction that HPR should localize to heterochromatin DNA regions prior to histone H3K9 methylation. To test this, we can use *SMT3* and *UBC9* (generated by Qianyi Yang) conditional strains and

transfect them with GFP-HP1. Accumulation of HPR to heterochromatic regions in *SMT3*- and *UBC9*-depleted cells would suggest SUMO does not target HPR to heterochromatic regions of the developing MAC. A second method to test the same would be to make use of Lys mutants of GFP-HP1 that are unable to be SUMOylated expressed in wild-type *Tetrahymena*. If HPR is immunopurified using the V85K mutant, the exact site of Smt3p attachment on HP1 may be determined using the approach described above (see section on *Identification of SUMOylation sites on protein substrates*). Lys mutants of HPR would also help in identifying the site of Smt3p attachment. Western blot analysis would confirm the formation of SUMO-modified species of HP1 (versus un-modified) if the target Lys is present in HP1. The slow-migrating band corresponding to the Smt3p-HPR would be absent in HP1 Lys mutants. Chromatin Immunoprecipitation (ChIP) assay will confirm whether SUMO-HPR is physically associated with heterochromatic regions versus unmodified HPR1. SUMOylation of HPR may also modulate interactions with other proteins and could include other CHROMO domain-containing proteins such as Pdd1p and Pdd3p, both required for DNA elimination in *Tetrahymena* (96, 97).

Another SUMO target of particular interest is the histone H2A variant, HTA3 that was identified as a substrate of SUMO in *Tetrahymena*, yeast and fruit fly. In yeast, SUMOylation H2A.Z (HTA3 in *Tetrahymena*, TTHERM_00143660) serves as a specific chromatin mark for chromosome relocation to nuclear periphery for DSB repair (145, 165). Excision of IES elements in *Tetrahymena*, begins with induction of DSBs that eliminate mic-limited sequences followed by re-ligation of adjacent MAC-

destined DNA. These events occur within DNA rearrangement foci in the developing macronucleus. Our model is that SUMO modification of HTA3 is required for its localization to DSBs in DNA rearrangement foci. Utilizing specific IES elements such as the M- or R-element is useful here as the excision of these IESs is well studied. GFP-tagged HTA3 along with the use of anti-Smt3p antibody can be used in immunofluorescence experiments to determine whether the two proteins co-localize in the anlagen during MAC development. In addition to aiding DSB relocation, SUMOylation of HTA3 may be responsible for recruiting proteins involved in DSB repair by causing conformational changes that promote assembly of downstream repair factors. Homologous sequences are absent at DSB break sites and the cell switches to a type of Non-Homologous End Joining (NHEJ) repair mechanism as has been described in *Paramecium* (130). It is possible that the same process is in effect in *Tetrahymena* in the ligation of MAC-destined DNA following resection of IES elements. ChIP assays utilizing epitope-tagged versions of HTA3 can be used to identify interactions with other proteins and establish whether HTA3 is associated with eliminated heterochromatic on retained MAC-destined DNA regions. This will be especially helpful as Ku proteins and XRCC4 of the NHEJ pathways are known SUMO substrates and SUMOylation has well-defined roles in DNA repair (128). Conditional mutants of the SUMOylation pathway can be used to test for decreased SUMOylation status of HTA3. The SUMOylation-deficient strain (ΔGG) can also be used to check reduction of histone SUMOylation. And RNAi-based approach

to reduce SUMOylation pathway protein levels can be effectively used to test whether cells accumulate DSBs in the anlagen as conjugation progresses.

In vitro validation of Smt3p Conjugates found through Global MS Analysis

The zinc-finger C2H2 protein and Glycine Cleavage System protein (TTHERM_00499390, here referred to as HPM1) are among the Smt3p proteins identified in our two-step purification procedure in *Tetrahymena* (Chapter 3) that were also identified in yeast or fruit fly. These proteins are of special interest because microarray expression data shows that transcript levels are upregulated for these proteins during conjugation, specifically during MAC development. This would suggest that these proteins are involved in the nuclear events of MAC development.

We can validate these protein substrates as Smt3p conjugation targets using a bacterial SUMOylation assay. For this *in vitro* assay, we will generate *E. coli* expressed enzymes of the SUMOylation pathway from *Tetrahymena thermophila* (Smt3p, Aos1p/Uba2p, and Ubc9p) as well as epitope-tagged candidate Smt3p substrate. Single lysine mutants of Smt3p protein substrates will be used to confirm the exact site of Smt3p attachment. The detection of Smt3p conjugates in this system is facilitated by the lack of an absolute requirement for E3-type ligases in SUMOylation as well as by the absence of Smt3p proteases that deconjugate Smt3p from the substrate (reviewed in (2)). While bacterial SUMOylation systems may not completely recapitulate the specificity of SUMOylation reaction *in vivo*, they have

been repeatedly validated as a useful approach for confirming SUMOylation targets (48, 166).

In vivo validation of Smt3p Conjugates found through Global MS Analysis

To clearly demonstrate that a particular protein (e.g. HPM1, HTA3) is a target of Smt3p, an *in vivo* approach is most helpful in addition to the *in vitro* method described above. Most commonly this approach involves immunoprecipitation of the epitope-tagged form of the target protein. The use of denaturing lysis and affinity purification conditions helps to ensure that they type of interaction between the target protein and Smt3p is covalent in nature. This also helps in removal of non-covalently bound proteins as well as in the inhibition of SUMO proteases. Presence of a band in the immunoblot (probed with antibodies against the epitope tag) will confirm that the protein target is indeed a substrate of Smt3p. If antibodies for the protein substrates are available, then we can expect two bands in the immunoblots – one corresponding to the unmodified form of the protein substrate and another slower migrating band that will correspond to the SUMO-modified form.

Non-ligatable Smt3p mutants serve as a good control for substrate validation

The inactivated forms of UBLs require maturation before participation in the conjugation reaction. UBL proteases cleave the terminal residues of UBLs to reveal a diglycine (GG) motif that is necessary for isopeptide bond formation with the

acceptor Lys residue on target proteins. The evolutionary conservation of the GG motif across UBLs would suggest an absolute requirement for the GG motif in attachment to protein substrates. We have generated a series of C-terminal mutants of Smt3p all expressing the same His6-3FLAG double epitope tag described in Chapter 3. Immunoblotting analyses have shown that the deletions of any or all of the GG motif (GΔG and ΔGG) prevents formation of Smt3p adducts. These mutant strains where Smt3p cannot be conjugated to target proteins are useful in the affinity purification procedures used for the proteomics studies as proteins that are carried through the two purifications are highly abundant proteins and can be safely removed from the tagged Smt3p dataset.

Undoubtedly more work will have to be undertaken to completely elucidate the protein substrates of Smt3p as well as the mechanisms by which Smt3p modification of these proteins regulates critical cellular pathways. For these purposes, *Tetrahymena* is a very useful model system as it exhibits nuclear dimorphism and allows for powerful forward and reverse genetics which will certainly provide insights into the essential nature of Smt3p. Studies such as ours are the first step in this approach and yield useful catalogue of proteins that are Smt3p substrates.

REFERENCES

REFERENCES

1. **Johnson ES**. 2004. Protein modification by SUMO. Annu. Rev. Biochem. **73**:355–82.

2. **Gareau JR, Lima CD**. 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. Nat. Rev. Mol. Cell Biol. **11**:861–71.

3. **Guo D, Li M, Zhang Y, Yang P, Eckenrode S, Hopkins D, Zheng W, Purohit S, Podolsky RH, Muir A, Wang J, Dong Z, Brusko T, Atkinson M, Pozzilli P, Zeidler A, Raffel LJ, Jacob CO, Park Y, Serrano-Rios M, Larrad MT, Zhang Z, Garchon HJ, Bach JF, Rotter JI, She JX, Wang CY**. 2004. A functional variant of SUMO4, a new I kappa B alpha modifier, is associated with type 1 diabetes. Nat Genet **36**:837–841.

4. **Azuma Y, Arnaoutov A, Dasso M**. 2003. SUMO-2/3 regulates topoisomerase II in mitosis. J. Cell Biol. **163**:477–87.

5. **Zhang X-D, Goeres J, Zhang H, Yen TJ, Porter ACG, Matunis MJ**. 2008. SUMO-2/3 modification and binding regulate the association of CENP-E with kinetochores and progression through mitosis. Mol. Cell **29**:729–41.

6. **Azuma Y, Tan SH, Cavenagh MM, Ainsztein AM, Saitoh H, Dasso M**. 2001. Expression and regulation of the mammalian SUMO-1 E1 enzyme. FASEB J **15**:1825–1827.

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

8. **Bayer P, Arndt A, Metzger S, Mahajan R, Melchior F, Jaenicke R, Becker J**. 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. J. Mol. Biol. **280**:275–86.

9. **Mossessova E, Lima CD**. 2000. Ulp1-SUMO Crystal Structure and Genetic Analysis Reveal Conserved Interactions and a Regulatory Element Essential for Cell Growth in Yeast. Mol. Cell **5**:865–876.

10. **Müller S, Hoege C, Pyrowolakis G, Jentsch S**. 2001. SUMO, ubiquitin's mysterious cousin. Nat. Rev. Mol. Cell Biol. **2**:202–10.

11. **Vertegaal ACO**. 2010. SUMO chains: polymeric signals. Biochem. Soc. Trans. **38**:46–9.

12. **Matic I, van Hagen M, Schimmel J, Macek B, Ogg SC, Tatham MH, Hay RT, Lamond AI, Mann M, Vertegaal ACO**. 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**:132–44.

13. **Bylebyl GR, Belichenko I, Johnson ES**. 2003. The SUMO isopeptidase Ulp2 prevents accumulation of SUMO chains in yeast. J. Biol. Chem. **278**:44113–20.

14. **Hickey CM, Wilson NR, Hochstrasser M**. 2012. Function and regulation of SUMO proteases. Nat Rev Mol Cell Biol **13**:755–766.

15. **Hochstrasser M**. 1996. Ubiquitin-dependent protein degradation. Annu. Rev. Genet. **30**:405–39.

16. **Seeler JS, Bischof O, Nacerddine K, Dejean A**. 2007. SUMO, the three Rs and cancer. Curr. Top. Microbiol. Immunol. **313**:49–71.

17. **Sarge KD, Park-Sarge O-K**. 2009. Sumoylation and human disease pathogenesis. Trends Biochem. Sci. **34**:200–5.

18. **Lee J-S, Thorgeirsson SS**. 2004. Genome-scale profiling of gene expression in hepatocellular carcinoma: classification, survival prediction, and identification of therapeutic targets. Gastroenterology **127**:S51–5.

19. **Kessler JD, Kahle KT, Sun T, Meerbrey KL, Schlabach MR, Schmitt EM, Skinner SO, Xu Q, Li MZ, Hartman ZC, Rao M, Yu P, Dominguez-Vidana R, Liang AC, Solimini NL, Bernardi RJ, Yu B, Hsu T, Golding I, Luo J, Osborne CK, Creighton CJ, Hilsenbeck SG, Schiff R, Shaw CA, Elledge SJ, Westbrook TF**. 2012. A SUMOylation-dependent transcriptional subprogram is required for Myc-driven tumorigenesis. Science **335**:348–53.

20. **Kim K Il, Baek SH**. 2006. SUMOylation code in cancer development and metastasis. Mol. Cells **22**:247–53.

21. **Wang L, Banerjee S**. 2004. Differential PIAS3 expression in human malignancy. Oncol. Rep. **11**:1319–24.

22. **Cheng J, Bawa T, Lee P, Gong L, Yeh ETH**. 2006. Role of desumoylation in the development of prostate cancer. Neoplasia **8**:667–76.

23. **Xu Y, Li J, Zuo Y, Deng J, Wang L-S, Chen G-Q**. 2011. SUMO-specific protease 1 regulates the in vitro and in vivo growth of colon cancer cells with the upregulated expression of CDK inhibitors. Cancer Lett. **309**:78–84.

24. **Dorval V, Fraser PE**. 2007. SUMO on the road to neurodegeneration. Biochim. Biophys. Acta **1773**:694–706.

25. **Gao C, Huang W, Kanasaki K, Xu Y**. 2014. The role of ubiquitination and sumoylation in diabetic nephropathy. Biomed Res. Int. **2014**:160692.

26. **Leavenworth JW, Ma X, Mo Y, Pauza ME**. 2009. SUMO conjugation contributes to immune deviation in nonobese diabetic mice by suppressing c-Maf transactivation of IL-4. J. Immunol. **183**:1110–9.

27. **Hwang KW, Won TJ, Kim H, Chun H-J, Chun T, Park Y**. 2011. Characterization of the regulatory roles of the SUMO. Diabetes. Metab. Res. Rev. **27**:854–61.

28. **Kim EY, Zhang Y, Beketaev I, Segura AM, Yu W, Xi Y, Chang J, Wang J**. 2015. SENP5, a SUMO isopeptidase, induces apoptosis and cardiomyopathy. J. Mol. Cell. Cardiol. **78C**:154–164.

29. **Gupta MK, Gulick J, Liu R, Wang X, Molkentin JD, Robbins J**. 2014. Sumo E2 enzyme UBC9 is required for efficient protein quality control in cardiomyocytes. Circ. Res. **115**:721–9.

30. **Flotho A, Melchior F**. 2013. Sumoylation: a regulatory protein modification in health and disease. Annu Rev Biochem **82**:357–385.

31. **Desterro JM, Thomson J, Hay RT**. 1997. Ubch9 conjugates SUMO but not ubiquitin. FEBS Lett **417**:297–300.

32. **Johnson ES, Schwienhorst I, Dohmen RJ, Blobel G**. 1997. The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer. EMBO J. **16**:5509–19.

33. **Dohmen RJ, Stappen R, McGrath JP, Forrova H, Kolarov J, Goffeau A, Varshavsky A**. 1995. An essential yeast gene encoding a homolog of ubiquitin-activating enzyme. J Biol Chem, 1995/07/28 ed. **270**:18099–18109.

34. **Desterro JM, Rodriguez MS, Kemp GD, Hay RT**. 1999. Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. J. Biol. Chem. **274**:10618–24.

35. **McGrath JP, Jentsch S, Varshavsky A**. 1991. UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme. EMBO J. **10**:227–36.

36. **Jin J, Li X, Gygi SP, Harper JW**. 2007. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature **447**:1135–8.

37. **Gong L, Li B, Millas S, Yeh ET**. 1999. Molecular cloning and characterization of human AOS1 and UBA2, components of the sentrin-activating enzyme complex. FEBS Lett. **448**:185–9.

38. **Shayeghi M, Doe CL, Tavassoli M, Watts FZ**. 1997. Characterisation of Schizosaccharomyces pombe rad31, a UBA-related gene required for DNA damage tolerance. Nucleic Acids Res. **25**:1162–9.

39. **Donaghue C, Bates H, Cotterill S**. 2001. Identification and characterisation of the Drosophila homologue of the yeast Uba2 gene. Biochim. Biophys. Acta **1518**:210–4.

40. **Miao W, Xiong J, Bowen J, Wang W, Liu Y, Braguinets O, Grigull J, Pearlman RE, Orias E, Gorovsky MA**. 2009. Microarray analyses of gene expression during the Tetrahymena thermophila life cycle. PLoS One **4**:e4429.

41. **Geiss-Friedlander R, Melchior F**. 2007. Concepts in sumoylation: a decade on. Nat. Rev. Mol. Cell Biol. **8**:947–56.

42. **Anckar J, Sistonen L**. 2007. SUMO: getting it on. Biochem. Soc. Trans. **35**:1409–13.

43. **Melchior F**. 2000. SUMO--nonclassical ubiquitin. Annu. Rev. Cell Dev. Biol. **16**:591–626.

44. **Elrouby N, 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**:17415–20.

45. **Denison C, Rudner AD, Gerber S a, Bakalarski CE, Moazed D, Gygi SP**. 2005. A proteomic strategy for gaining insights into protein sumoylation in yeast. Mol. Cell. Proteomics **4**:246–54.

46. **Tammsalu T, Matic I, Jaffray EG, Ibrahim AFM, Tatham MH, Hay RT**. 2014. Proteome-wide identification of SUMO2 modification sites. Sci. Signal. **7**:rs2.

47. **Golebiowski F, Matic I, Tatham MH, Cole C, Yin Y, Nakamura A, Cox J, Barton GJ, Mann M, Hay RT**. 2009. System-wide changes to SUMO modifications in response to heat shock. Sci. Signal. **2**:ra24.

48. **Nie M, Xie Y, Loo JA, Courey AJ**. 2009. Genetic and proteomic evidence for roles of Drosophila SUMO in cell cycle control, Ras signaling, and early pattern formation. PLoS One **4**:e5905.

49. **Wilkinson KD**. 1997. Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. FASEB J. **11**:1245–56.

50. **Eletr ZM, Wilkinson KD**. 2014. Regulation of proteolysis by human deubiquitinating enzymes. Biochim. Biophys. Acta **1843**:114–28.

51. **Li S-J, Hochstrasser M**. 1999. A new protease required for cell-cycle progression in yeast. Nature **398**:246–251.

52. **Li SJ, Hochstrasser M**. 2000. The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. Mol. Cell. Biol. **20**:2367–77.

53. **Miller MJ, Scalf M, Rytz TC, Hubler SL, Smith LM, Vierstra RD**. 2013. Quantitative proteomics reveals factors regulating RNA biology as dynamic targets of stress-induced SUMOylation in Arabidopsis. Mol. Cell. Proteomics **12**:449–63.

54. **Bossis G, Melchior F**. 2006. Regulation of SUMOylation by reversible oxidation of SUMO conjugating enzymes. Mol. Cell **21**:349–57.

55. **Boggio R, Colombo R, Hay RT, Draetta GF, Chiocca S**. 2004. A mechanism for inhibiting the SUMO pathway. Mol. Cell **16**:549–61.

56. **Boggio R, Passafaro A, Chiocca S**. 2007. Targeting SUMO E1 to ubiquitin ligases: a viral strategy to counteract sumoylation. J. Biol. Chem. **282**:15376–82.

57. **Ribet D, Hamon M, Gouin E, Nahori M-A, Impens F, Neyret-Kahn H, Gevaert K, Vandekerckhove J, Dejean A, Cossart P**. 2010. Listeria monocytogenes impairs SUMOylation for efficient infection. Nature **464**:1192–5.

58. **Makhnevych T, Ptak C, Lusk CP, Aitchison JD, Wozniak RW**. 2007. The role of karyopherins in the regulated sumoylation of septins. J. Cell Biol. **177**:39–49.

59. **Sydorskyy Y, Srikumar T, Jeram SM, Wheaton S, Vizeacoumar FJ, Makhnevych T, Chong YT, Gingras A-C, Raught B**. 2010. A novel mechanism for SUMO system control: regulated Ulp1 nucleolar sequestration. Mol. Cell. Biol. **30**:4452–62.

60. **Pichler A, Knipscheer P, Saitoh H, Sixma TK, Melchior F**. 2004. The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. Nat. Struct. Mol. Biol. **11**:984–91.

61. **Ulrich HD**. 2005. Mutual interactions between the SUMO and ubiquitin systems: a plea of no contest. Trends Cell Biol. **15**:525–32.

62. **Denuc A, Marfany G**. 2010. SUMO and ubiquitin paths converge. Biochem. Soc. Trans. **38**:34–9.

63. **Desterro JM, Rodriguez MS, Hay RT**. 1998. SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. Mol Cell **2**:233–239.

64. **Dohmen RJ**. 2004. SUMO protein modification. Biochim. Biophys. Acta **1695**:113–31.

65. **Prudden J, Pebernard S, Raffa G, Slavin DA, Perry JJP, Tainer JA, McGowan CH, Boddy MN**. 2007. SUMO-targeted ubiquitin ligases in genome stability. EMBO J. **26**:4089–101.

66. **Yang X-J, Grégoire S**. 2006. A recurrent phospho-sumoyl switch in transcriptional repression and beyond. Mol. Cell **23**:779–86.

67. **Ullmann R, Chien CD, Avantaggiati ML, Muller S**. 2012. An acetylation switch regulates SUMO-dependent protein interaction networks. Mol. Cell **46**:759–70.

68. **Stankovic-Valentin N, Deltour S, Seeler J, Pinte S, Vergoten G, Guérardel C, Dejean A, Leprince D**. 2007. An acetylation/deacetylation-SUMOylation switch through a phylogenetically conserved psiKXEP motif in the tumor suppressor HIC1 regulates transcriptional repression activity. Mol. Cell. Biol. **27**:2661–75.

69. **Salinas S, Briançon-Marjollet A, Bossis G, Lopez M-A, Piechaczyk M, Jariel-Encontre I, Debant A, Hipskind RA**. 2004. SUMOylation regulates nucleo-cytoplasmic shuttling of Elk-1. J. Cell Biol. **165**:767–73.

70. **Stielow B, Krüger I, Diezko R, Finkernagel F, Gillemans N, Kong-a-San J, Philipsen S, Suske G**. 2010. Epigenetic silencing of spermatocyte-specific and neuronal genes by SUMO modification of the transcription factor Sp3. PLoS Genet. **6**:e1001203.

71. **Shiio Y, Eisenman RN**. 2003. Histone sumoylation is associated with transcriptional repression. Proc. Natl. Acad. Sci. U. S. A. **100**:13225–30.

72. **Baba D, Maita N, Jee J-G, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M**. 2005. Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. Nature **435**:979–82.

73. **Hardeland U, Steinacher R, Jiricny J, Schar P**. 2002. Modification of the human thymine-DNA glycosylase by ubiquitin-like proteins facilitates enzymatic turnover. EMBO J **21**:1456–1464.

74. **Steinacher R, Schär P**. 2005. Functionality of human thymine DNA glycosylase requires SUMO-regulated changes in protein conformation. Curr. Biol. **15**:616–23.

75. **Galanty Y, Belotserkovskaya R, Coates J, Polo S, Miller KM, Jackson SP**. 2009. Mammalian SUMO E3-ligases PIAS1 and PIAS4 promote responses to DNA double-strand breaks. Nature **462**:935–9.

76. **Morris JR, Boutell C, Keppler M, Densham R, Weekes D, Alamshah A, Butler L, Galanty Y, Pangon L, Kiuchi T, Ng T, Solomon E**. 2009. The SUMO modification pathway is involved in the BRCA1 response to genotoxic stress. Nature **462**:886–90.

77. **Tanaka K, Nishide J, Okazaki K, Kato H, Niwa O, Nakagawa T, Matsuda H, Kawamukai M, Murakami Y**. 1999. Characterization of a fission yeast SUMO-1 homologue, pmt3p, required for multiple nuclear events, including the control of telomere length and chromosome segregation. Mol. Cell. Biol. **19**:8660–72.

78. **Schwienhorst I, Johnson ES, Dohmen RJ**. 2000. SUMO conjugation and deconjugation. Mol. Gen. Genet. **263**:771–86.

79. **Dieckhoff P, Bolte M, Sancak Y, Braus GH, Irniger S**. 2004. Smt3/SUMO and Ubc9 are required for efficient APC/C-mediated proteolysis in budding yeast. Mol Microbiol, 2004/02/26 ed. **51**:1375–1387.

80. **Zhong S, Müller S, Ronchetti S, Freemont PS, Dejean A, Pandolfi PP**. 2000. Role of SUMO-1-modified PML in nuclear body formation. Blood **95**:2748–52.

81. **Nacerddine K, Lehembre F, Bhaumik M, Artus J, Cohen-Tannoudji M, Babinet C, Pandolfi PP, Dejean A**. 2005. The SUMO pathway is essential for nuclear integrity and chromosome segregation in mice. Dev. Cell **9**:769–79.

82. **Dasso M**. 2008. Emerging roles of the SUMO pathway in mitosis. Cell Div. **3**:5.

83. **Lee M-T, Bachant J**. 2009. SUMO modification of DNA topoisomerase II: trying to get a CENse of it all. DNA Repair (Amst). **8**:557–68.

84. **De Carvalho CE, Colaiácovo MP**. 2006. SUMO-mediated regulation of synaptonemal complex formation during meiosis. Genes Dev. **20**:1986–92.

85. **Cheng C-H, Lo Y-H, Liang S-S, Ti S-C, Lin F-M, Yeh C-H, Huang H-Y, Wang T-F**. 2006. SUMO modifications control assembly of synaptonemal complex and polycomplex in meiosis of Saccharomyces cerevisiae. Genes Dev. **20**:2067–81.

86. **La Salle S, Sun F, Zhang X-D, Matunis MJ, Handel MA**. 2008. Developmental control of sumoylation pathway proteins in mouse male germ cells. Dev. Biol. **321**:227–37.

87. **Vigodner M, Ishikawa T, Schlegel PN, Morris PL**. 2006. SUMO-1, human male germ cell development, and the androgen receptor in the testis of men with normal and abnormal spermatogenesis. Am. J. Physiol. Endocrinol. Metab. **290**:E1022–33.

88. **Smith M, Turki-Judeh W, Courey AJ**. 2012. SUMOylation in Drosophila Development. Biomolecules **2**:331–49.

89. **Lomelí H, Vázquez M**. 2011. Emerging roles of the SUMO pathway in development. Cell. Mol. Life Sci. **68**:4045–64.

90. **Saitoh H, Hinchey J**. 2000. Functional heterogeneity of small ubiquitin-related protein modifiers SUMO-1 versus SUMO-2/3. J Biol Chem **275**:6252–6258.

91. **Hay RT**. 2005. SUMO: a history of modification. Mol. Cell **18**:1–12.

92. **Wilson VG, Heaton PR**. 2008. Ubiquitin proteolytic system: focus on SUMO. Expert Rev. Proteomics **5**:121–35.

93. **Filosa G, Barabino SML, Bachi A**. 2013. Proteomics strategies to identify SUMO targets and acceptor sites: a survey of RNA-binding proteins SUMOylation. Neuromolecular Med. **15**:661–76.

94. **Hietakangas V, Ahlskog JK, Jakobsson AM, Hellesuo M, Sahlberg NM, Holmberg CI, Mikhailov A, Palvimo JJ, Pirkkala L, Sistonen L**. 2003. Phosphorylation of serine 303 is a prerequisite for the stress-inducible SUMO modification of heat shock factor 1. Mol. Cell. Biol. **23**:2953–68.

95. **Müller S, Miller WH, Dejean A**. 1998. Trivalent antimonials induce degradation of the PML-RAR oncoprotein and reorganization of the promyelocytic leukemia nuclear bodies in acute promyelocytic leukemia NB4 cells. Blood **92**:4308–16.

96. **Chalker DL**. 2008. Dynamic nuclear reorganization during genome remodeling of Tetrahymena. Biochim. Biophys. Acta **1783**:2130–6.

97. **Bracht JR, Fang W, Goldman AD, Dolzhenko E, Stein EM, Landweber LF**. 2013. Genomes on the edge: programmed genome instability in ciliates. Cell **152**:406–16.

98. **Fan Q, Yao M**. 1996. New telomere formation coupled with site-specific chromosome breakage in Tetrahymena thermophila. Mol. Cell. Biol. **16**:1267–74.

99. **Yu GL, Bradley JD, Attardi LD, Blackburn EH**. 1990. In vivo alteration of telomere sequences and senescence caused by mutated Tetrahymena telomerase RNAs. Nature **344**:126–32.

100. **Altschuler MI, Yao M-C**. 1985. Macronuclear DNA of Tetrahymena thermophila exists as defined subchromosomal-sized molecules. Nucleic Acids Res. **13**:5817–5831.

101. **Taverna SD, Coyne RS, Allis CD**. 2002. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. Cell **110**:701–11.

102. **Yao M-C, Chao J-L**. 2005. RNA-guided DNA deletion in Tetrahymena: an RNAi-based mechanism for programmed genome rearrangements. Annu. Rev. Genet. **39**:537–59.

103. **Schoeberl UE, Mochizuki K**. 2011. Keeping the soma free of transposons: programmed DNA elimination in ciliates. J. Biol. Chem. **286**:37045–52.

104. **Coyne RS, Thiagarajan M, Jones KM, Wortman JR, Tallon LJ, Haas BJ, Cassidy-Hanley DM, Wiley EA, Smith JJ, Collins K, Lee SR, Couvillion MT, Liu Y, Garg J, Pearlman RE, Hamilton EP, Orias E, Eisen JA, Methé BA**. 2008. Refined annotation and assembly of the Tetrahymena thermophila genome sequence through EST analysis, comparative genomic hybridization, and targeted gap closure. BMC Genomics **9**:562.

105. **Robinson R**. 2006. Ciliate genome sequence reveals unique features of a model eukaryote. PLoS Biol. **4**:e304.

106. **Eisen JA, Coyne RS, Wu M, Wu D, Thiagarajan M, Wortman JR, Badger JH, Ren Q, Amedeo P, Jones KM, Tallon LJ, Delcher AL, Salzberg SL, Silva JC, Haas BJ, Majoros WH, Farzad M, Carlton JM, Smith RK, Garg J, Pearlman RE, Karrer KM, Sun L, Manning G, Elde NC, Turkewitz AP, Asai DJ, Wilkes DE, Wang Y, Cai H, Collins K, Stewart BA, Lee SR, Wilamowska K, Weinberg Z, Ruzzo WL, Wloga D, Gaertig J, Frankel J, Tsao C-C, Gorovsky MA, Keeling PJ, Waller RF, Patron NJ, Cherry JM, Stover NA, Krieger CJ, del Toro C, Ryder HF, Williamson SC, Barbeau RA, Hamilton EP, Orias E**. 2006. Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS Biol. **4**:e286.

107. **Duharcourt S, Yao M-C**. 2002. Role of Histone Deacetylation in Developmentally Programmed DNA Rearrangements in Tetrahymena thermophila. Eukaryot. Cell **1**:293–303.

108. **Coyne RS, Chalker DL, Yao MC**. 1996. Genome downsizing during ciliate development: nuclear division of labor through chromosome restructuring. Annu. Rev. Genet. **30**:557–78.

109. **Wohlschlegel JA, Johnson ES, Reed SI, Yates JR**. 2004. Global analysis of protein sumoylation in Saccharomyces cerevisiae. J. Biol. Chem. **279**:45662–8.

110. **Nowak M, Hammerschmidt M**. 2006. Ubc9 regulates mitosis and cell survival during zebrafish development. Mol. Biol. Cell **17**:5324–36.

111. **Matsuda A, Forney JD**. 2006. The SUMO pathway is developmentally regulated and required for programmed DNA elimination in Paramecium tetraurelia. Eukaryot. Cell **5**:806–15.

112. **Nasir AM, Yang Q, Chalker DL, Forney JD**. 2014. SUMOylation is Developmentally Regulated and Required for Cell Pairing during Conjugation in Tetrahymena. Eukaryot. Cell **14**:170–81.

113. **Shang Y, Song X, Bowen J, Corstanje R, Gao Y, Gaertig J, Gorovsky MA**. 2002. A robust inducible-repressible promoter greatly facilitates gene knockouts, conditional expression, and overexpression of homologous and heterologous genes in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. **99**:3734–9.

114. **Bruns PJ, Cassidy-Hanley D**. 1999. Methods in Cell Biology Volume 62. Elsevier.

115. **Hai B, Gaertig J, Gorovsky MA**. 2000. Knockout heterokaryons enable facile mutagenic analysis of essential genes in Tetrahymena. Methods Cell Biol. **62**:513–31.

116. **Matsuda A, Shieh AW-Y, Chalker DL, Forney JD**. 2010. The conjugation-specific Die5 protein is required for development of the somatic nucleus in both Paramecium and Tetrahymena. Eukaryot. Cell **9**:1087–99.

117. **Gaertig J, Thatcher TH, Gu L, Gorovsky MA**. 1994. Electroporation-mediated replacement of a positively and negatively selectable beta-tubulin gene in Tetrahymena thermophila. Proc. Natl. Acad. Sci. U. S. A. **91**:4549–53.

118. **Stover NA, Krieger CJ, Binkley G, Dong Q, Fisk DG, Nash R, Sethuraman A, Weng S, Cherry JM**. 2006. Tetrahymena Genome Database (TGD): a new genomic resource for Tetrahymena thermophila research. Nucleic Acids Res. **34**:D500–3.

119. **Jones D, Crowe E, Stevens TA, Candido EPM**. 2002. Functional and phylogenetic analysis of the ubiquitylation system in Caenorhabditis elegans: ubiquitin-conjugating enzymes, ubiquitin-activating enzymes, and ubiquitin-like proteins. Genome Biol. **3**:RESEARCH0002.

120. **Saracco SA, Miller MJ, Kurepa J, Vierstra RD**. 2007. Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. Plant Physiol. **145**:119–34.

121. **Ho JC, Warr NJ, Shimizu H, Watts FZ**. 2001. SUMO modification of Rad22, the Schizosaccharomyces pombe homologue of the recombination protein Rad52. Nucleic Acids Res. **29**:4179–86.

122. **Hoege C, Pfander B, Moldovan G-L, Pyrowolakis G, Jentsch S**. 2002. RAD6-dependent DNA repair is linked to modification of PCNA by ubiquitin and SUMO. Nature **419**:135–41.

123. **Maeda D, Seki M, Onoda F, Branzei D, Kawabe Y-I, Enomoto T**. 2004. Ubc9 is required for damage-tolerance and damage-induced interchromosomal homologous recombination in S. cerevisiae. DNA Repair (Amst). **3**:335–41.

124. **Frankenberg-Schwager M, Kirchermeier D, Greif G, Baer K, Becker M, Frankenberg D**. 2005. Cisplatin-mediated DNA double-strand breaks in replicating but not in quiescent cells of the yeast Saccharomyces cerevisiae. Toxicology **212**:175–84.

125. **Cunha D, Cunha R, Côrte-Real M, Chaves SR**. 2013. Cisplatin-induced cell death in Saccharomyces cerevisiae is programmed and rescued by proteasome inhibition. DNA Repair (Amst). **12**:444–9.

126. **Ormerod MG, Orr RM, Peacock JH**. 1994. The role of apoptosis in cell killing by cisplatin: a flow cytometric study. Br. J. Cancer **69**:93–100.

127. **Deyrieux AF, Rosas-Acosta G, Ozbun MA, Wilson VG**. 2007. Sumoylation dynamics during keratinocyte differentiation. J. Cell Sci. **120**:125–36.

128. **Yurchenko V, Xue Z, Sadofsky MJ**. 2006. SUMO modification of human XRCC4 regulates its localization and function in DNA double-strand break repair. Mol. Cell. Biol. **26**:1786–94.

129. **Jahn CL, Klobutcher LA**. 2002. Genome remodeling in ciliated protozoa. Annu. Rev. Microbiol. **56**:489–520.

130. **Kapusta A, Matsuda A, Marmignon A, Ku M, Silve A, Meyer E, Forney JD, Malinsky S, Bétermier M**. 2011. Highly precise and developmentally programmed genome assembly in Paramecium requires ligase IV-dependent end joining. PLoS Genet. **7**:e1002049.

131. **Wang Y, Dohlman HG**. 2006. Pheromone-regulated sumoylation of transcription factors that mediate the invasive to mating developmental switch in yeast. J. Biol. Chem. **281**:1964–9.

132. **Nixon CE, Wilcox AJ, Laney JD**. 2010. Degradation of the Saccharomyces cerevisiae mating-type regulator alpha1: genetic dissection of cis-determinants and trans-acting pathways. Genetics **185**:497–511.

133. **Cervantes MD, Hamilton EP, Xiong J, Lawson MJ, Yuan D, Hadjithomas M, Miao W, Orias E**. 2013. Selecting one of several mating types through gene segment joining and deletion in Tetrahymena thermophila. PLoS Biol. **11**:e1001518.

134. **Wykoff DD, O'Shea EK**. 2005. Identification of sumoylated proteins by systematic immunoprecipitation of the budding yeast proteome. Mol. Cell. Proteomics **4**:73–83.

135. **Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP**. 2003. A proteomics approach to understanding protein ubiquitination. Nat. Biotechnol. **21**:921–6.

136. **Maison C, Bailly D, Roche D, Montes de Oca R, Probst A V, Vassias I, Dingli F, Lombard B, Loew D, Quivy J-P, Almouzni G**. 2011. SUMOylation promotes de novo targeting of HP1α to pericentric heterochromatin. Nat. Genet. **43**:220–7.

137. **Bonacci T, Audebert S, Camoin L, Baudelet E, Bidaut G, Garcia M, Witzel I-I, Perkins ND, Borg J-P, Iovanna J-L, Soubeyran P**. 2014. Identification of new mechanisms of cellular response to chemotherapy by tracking changes in post-translational modifications by ubiquitin and ubiquitin-like proteins. J. Proteome Res. **13**:2478–94.

138. **Galisson F, Mahrouche L, Courcelles M, Bonneil E, Meloche S, Chelbi-Alix MK, Thibault P**. 2011. A novel proteomics approach to identify SUMOylated proteins and their modification sites in human cells. Mol. Cell. Proteomics **10**:M110.004796.

139. **Cooper HJ, Tatham MH, Jaffray E, Heath JK, Lam TT, Marshall AG, Hay RT**. 2005. Fourier transform ion cyclotron resonance mass spectrometry for the analysis of small ubiquitin-like modifier (SUMO) modification: identification of lysines in RanBP2 and SUMO targeted for modification during the E3 autoSUMOylation reaction. Anal. Chem. **77**:6310–9.

140. **Knipscheer P, Flotho A, Klug H, Olsen J V, van Dijk WJ, Fish A, Johnson ES, Mann M, Sixma TK, Pichler A**. 2008. Ubc9 sumoylation regulates SUMO target discrimination. Mol. Cell **31**:371–82.

141. **Sung M-K, Lim G, Yi D-G, Chang YJ, Yang E Bin, Lee K, Huh W-K**. 2013. Genome-wide bimolecular fluorescence complementation analysis of SUMO interactome in yeast. Genome Res. **23**:736–46.

142. **Zhou W, Ryan JJ, Zhou H**. 2004. Global analyses of sumoylated proteins in Saccharomyces cerevisiae. Induction of protein sumoylation by cellular stresses. J. Biol. Chem. **279**:32262–8.

143. **Chen J, Shi X, Padmanabhan R, Wang Q, Wu Z, Stevenson SC, Hild M, Garza D, Li H**. 2008. Identification of novel modulators of mitochondrial function by a genome-wide RNAi screen in Drosophila melanogaster. Genome Res. **18**:123–36.

144. **Ullah H, Scappini EL, Moon AF, Williams LV, Armstrong DL, Pedersen LC**. 2008. Structure of a signal transduction regulator, RACK1, from Arabidopsis thaliana. Protein Sci. **17**:1771–80.

145. **Kalocsay M, Hiller NJ, Jentsch S**. 2009. Chromosome-wide Rad51 spreading and SUMO-H2A.Z-dependent chromosome fixation in response to a persistent DNA double-strand break. Mol. Cell **33**:335–43.

146. **El McHichi B, Regad T, Maroui M-A, Rodriguez MS, Aminev A, Gerbaud S, Escriou N, Dianoux L, Chelbi-Alix MK**. 2010. SUMOylation promotes PML degradation during encephalomyocarditis virus infection. J. Virol. **84**:11634–45.

147. **Leach MD, Stead DA, Argo E, Brown AJP**. 2011. Identification of sumoylation targets, combined with inactivation of SMT3, reveals the impact of sumoylation upon growth, morphology, and stress resistance in the pathogen Candida albicans. Mol. Biol. Cell **22**:687–702.

148. **Park HC, Choi W, Park HJ, Cheong MS, Koo YD, Shin G, Chung WS, Kim W-Y, Kim MG, Bressan RA, Bohnert HJ, Lee SY, Yun D-J**. 2011. Identification and molecular properties of SUMO-binding proteins in Arabidopsis. Mol. Cells **32**:143–51.

149. **Panse VG, Hardeland U, Werner T, Kuster B, Hurt E**. 2004. A proteome-wide approach identifies sumoylated substrate proteins in yeast. J. Biol. Chem. **279**:41346–51.

150. **Klug H, Xaver M, Chaugule VK, Koidl S, Mittler G, Klein F, Pichler A**. 2013. Ubc9 sumoylation controls SUMO chain formation and meiotic synapsis in Saccharomyces cerevisiae. Mol. Cell **50**:625–36.

151. **Tatham MH, Jaffray E, Vaughan OA, Desterro JM, Botting CH, Naismith JH, Hay RT**. 2001. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. J. Biol. Chem. **276**:35368–74.

152. **Zhao Q, Xie Y, Zheng Y, Jiang S, Liu W, Mu W, Liu Z, Zhao Y, Xue Y, Ren J**. 2014. GPS-SUMO: a tool for the prediction of sumoylation sites and SUMO-interaction motifs. Nucleic Acids Res. **42**:W325–30.

153. **Saitoh H, Sparrow DB, Shiomi T, Pu RT, Nishimoto T, Mohun TJ, Dasso M**. 1998. Ubc9p and the conjugation of SUMO-1 to RanGAP1 and RanBP2. Curr. Biol. **8**:121–4.

154. **Gill G**. 2004. SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? Genes Dev. **18**:2046–59.

155. **Jackson SP, Durocher D**. 2013. Regulation of DNA damage responses by ubiquitin and SUMO. Mol. Cell **49**:795–807.

156. **Becker J, Barysch S V, Karaca S, Dittner C, Hsiao H-H, Berriel Diaz M, Herzig S, Urlaub H, Melchior F**. 2013. Detecting endogenous SUMO targets in mammalian cells and tissues. Nat. Struct. Mol. Biol. **20**:525–31.

157. **Panse VG, Kressler D, Pauli A, Petfalski E, Gnädig M, Tollervey D, Hurt E**. 2006. Formation and nuclear export of preribosomes are functionally linked to the small-ubiquitin-related modifier pathway. Traffic **7**:1311–21.

158. **Finley D, Bartel B, Varshavsky A**. 1989. The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. Nature **338**:394–401.

159. **Wen D, Xu Z, Xia L, Liu X, Tu Y, Lei H, Wang W, Wang T, Song L, Ma C, Xu H, Zhu W, Chen G, Wu Y**. 2014. Important role of SUMOylation of Spliceosome factors in prostate cancer cells. J. Proteome Res. **13**:3571–82.

160. **Wilson V**. 2009. SUMO Regulation of Cellular ProcessesIX. Springer Netherlands, Dordrecht.

161. **Wilson VG, Rosas-Acosta G**. 2005. Wrestling with SUMO in a new arena. Sci. STKE **2005**:pe32.

162. **Martin S, Wilkinson KA, Nishimune A, Henley JM**. 2007. Emerging extranuclear roles of protein SUMOylation in neuronal function and dysfunction. Nat. Rev. Neurosci. **8**:948–59.

163. **Anderson DB, Wilkinson KA, Henley JM**. 2009. Protein SUMOylation in neuropathological conditions. Drug News Perspect. **22**:255–65.

164. **Hannich JT, Lewis A, Kroetz MB, Li S-J, Heide H, Emili A, Hochstrasser M**. 2005. Defining the SUMO-modified proteome by multiple approaches in Saccharomyces cerevisiae. J. Biol. Chem. **280**:4102–10.

165. **Sevilla A, Binda O**. 2014. Post-translational modifications of the histone variant H2AZ. Stem Cell Res. **12**:289–95.

166. **Flotho A, Werner A, Winter T, Frank AS, Ehret H, Melchior F**. 2012. Recombinant reconstitution of sumoylation reactions in vitro. Methods Mol. Biol. **832**:93–110.

APPENDIX

APPENDIX

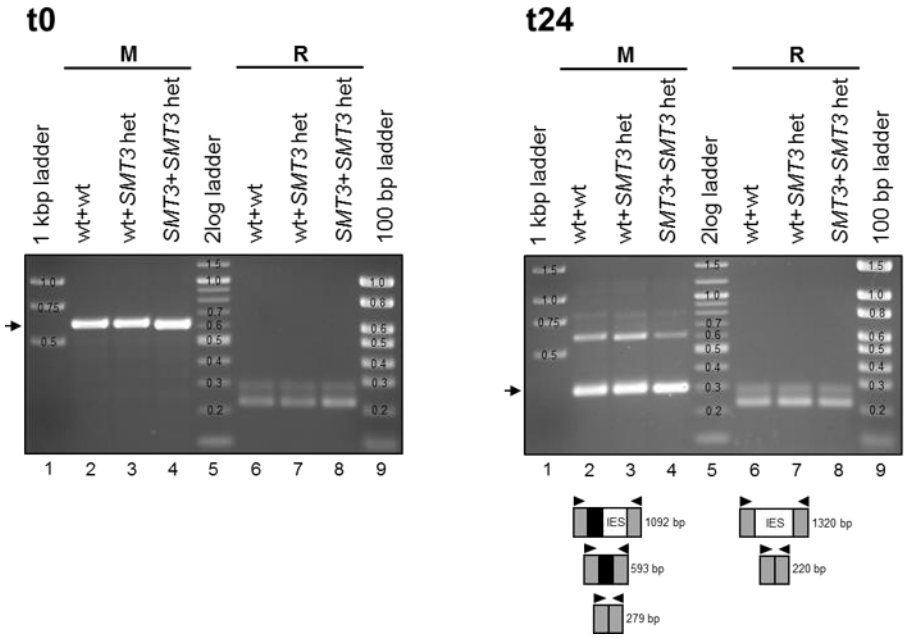

Figure 15. *SMT3 KO* heterokaryons undergo normal DNA elimination events during conjugation. Wild type (wt) and *SMT3 KO* heterokaryons (*SMT3het*) were mixed in three matings: wt × wt, wt × *SMT3het*, *SMT3het* × *SMT3het*. Cell lysates for PCR were prepared at mixing (t0). At 2 hours, >500 mated pairs were hand isolated into drops for each mating. After 24 hours (t24) when conjugation was complete the progeny were lysed and used as template in PCR reactions. The assay utilizes oligonucleotide primers that are able to amplify the unrearranged and rearranged forms of two IES elements (M and R). Comparison of PCR-amplified products on ethidium-bromide stained agarose gels, allowed visualization of the IES excision products. Only the longer 593 bp unrearranged form of the M element was found in the cells at t0, but the shorter 279 bp rearranged M element was the major product from progeny cells at t24 hours. This confirms that normal M element rearrangement occurred in progeny of SMT3 KO heterokaryons. There is only one R element product in the MAC but the amplified products are the same at t0 and t24 as expected in for normal excision. Schematic of IES products following arrangement are shown below.

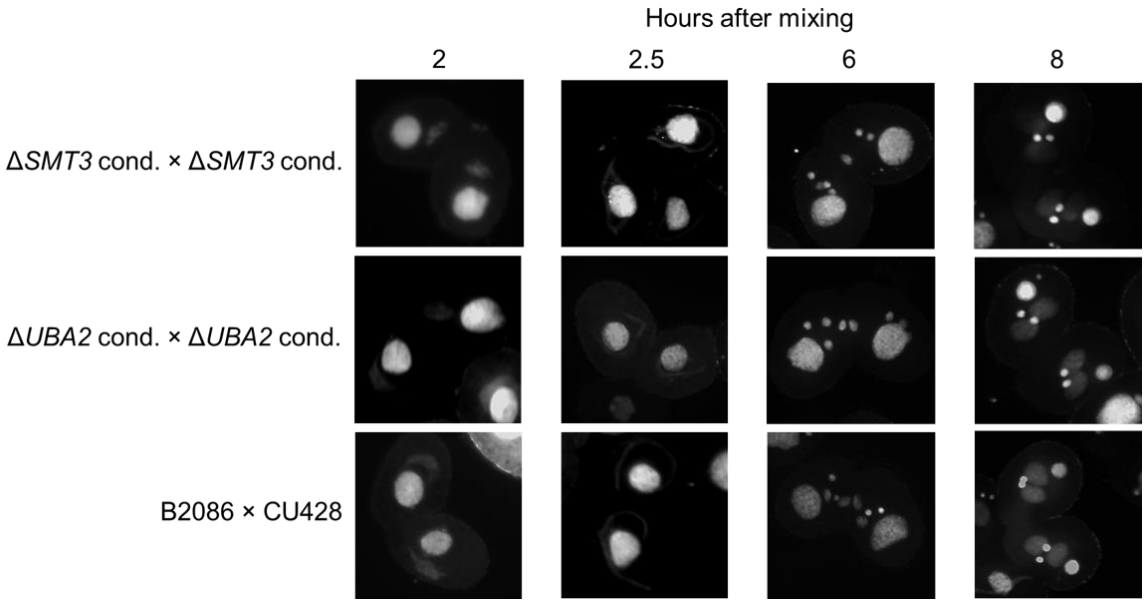

Figure 16. Induction of the *MTT1* promoter with cadmium allows *SMT3* and *UBA2* conditional mutants to conjugate normally. *SMT3* and *UBA2* conditional strains grown and starved in the absence of cadmium are unable to form pairs when mixed together to induce conjugation. When $CdCl_2$ (0.1 µg/mL) is added to these cells prior to mixing, these cells are able to form pairs and undergo normal conjugation. These mated cells were fixed and then stained with DAPI (to stain for nuclei). The time-points shown here are at 2-2.5 hours (crescent micronuclei), 6 hours (pronuclei formation) and 8 hours (MAC Development). Wild-type cells induced with the same amount of cadmium were used as control here.

Table 6. List of oligos used in this study.

Purpose	Name	Sequence (5'>3')
GFP constructs	bzSUMO+1F	CACCAATAAAATGACTGATTAAAACGCTAACGCT
	bzSUMO+930R	GATATCGAAAGAGCCACCAACTTGTTC
	bzUBA2+1F	CACCAAGGATATGAGTTTAGGAAGAATAAATC
	bzUBA2+2262R	GATATCCACTTTTAACTTTTTGTTGTCT
Knockout cassette	SMT3 5'flank F	GTCACTCGAGGCAGTTTGTCTTTTATCCATTT
	SMT3 5'flank R	GCCGAGATCTCTTCAAATATTTATTGTTCGAC
	SMT3 3'flank F	ATCGGGATCCCTTCAAAATTTAGTTGATTGTGATAACA
	SMT3 3'flank R	ATTAGCGGCCGCATCTCAAATAAGTCTAAAT
	UBA2 5'flank F	ATCGCTCGAGAGTACTCGACGGATCTCATAAA
	UBA2 5'flank R	CGGCGGATCCACTCATATCCTTATCAATTAAA
	UBA2 3'flank F	ACGTGGATCCAAAGCTGTAGATTTTAGTTAAA
	UBA2 3'flank R	AACAGCGGCCGCTAGCTTATTAATTCTTCTA
PCR to confirm knockout lines	SMT3WT upstream F	GATTGTTTGATGCTACATTCCTTC
	SMT3WT upstream R	TGCTAAAGACGGTTGGCTCT
	UBA2WT upstream F	TTTGCTTGTTGTTTGGTTTGTTT
	UBA2WT upstream R	CCAATGCCTCCTACACCAAT
	MTTp 1940R	TTTGCTAACCATAGCCAAAAT
RT-PCR assay of conditional line	SMT3WT 3'UTR F	ACTGATTAAAACGCTAACGCT
	SMT3WT 3'UTR R	ATTAGTTTATAAGCAAGCACATAC

	BTU1 3’UTR R	GTTGGTTTAGCTGACCGATTC
	NEO3 3’ UTR F	ACCGCTATCAGGACATA
	NEO3 3’ UTR R	GCAGACAAATTTTTAAGAGC

Table 7. List of identified Smt3p protein substrates from vegetative *Tetrahymena.*

TGD ID	Gene Name	Description	Independent Identified
TTHERM_00410130	*SMT3*	Small ubiquitin-related modifier (SUMO), conjugated to lysine residues on target proteins, post-translational modification.	5
TTHERM_00641240	*GPM1*	Phosphoglycerate mutase.	4
TTHERM_00585260	*ATP2*	ATP synthase beta chain, mitochondrial precursor, putative.	2
TTHERM_00339620	*RPL40*	Ubiquitin-60S ribosomal protein L40.	5
TTHERM_00346510		Ubiquitin family protein.	
TTHERM_00346560		Ubiquitin family protein.	3
TTHERM_00085200		Ubiquitin family protein.	2
TTHERM_01014750		dnaK protein belongs to the heat shock protein 70 family.	
TTHERM_00245410		Heterochromatin protein 1, putative, Chromo (CHRromatin Organisation MOdifier) domain.	5
TTHERM_00529790	*CIT1*	Citrate synthase; bifunctional 14-nm filament-forming protein; structural protein involved in oral morphogenesis and pronuclear behavior during conjugation; citrate synthase activity decreased by polymerization and dephosphorylation.	2
TTHERM_00196370	*HSP60*	HSP60; TCP-1/cpn60 chaperonin family protein.	
TTHERM_00399360		Hypothetical protein.	3
TTHERM_00140850		Hypothetical protein.	5
TTHERM_00185230		Triosephosphate isomerase.	2
TTHERM_00028780		Hypothetical protein.	
TTHERM_00046480	*ENO2*	ENOlase; enolase family protein.	

TTHERM_00105110	*HSP70*	Putative 70-kDa heat-shock protein (hsp70); two forms of hsp70 associate with tubulin, mitochondrial hsp70 associates with 14-nm filament protein (14FP), hsp70 proteins have evolutionarily conserved roles in folding nascent polypeptides.	
TTHERM_00125640	*SSA3*	HSP70a paralog; dnaK protein.	
TTHERM_00558440	*SSA5*	HSP70a paralog; dnaK protein.	
TTHERM_00143660	*HTA3*	Histone H2A (variant H2A.Z); essential protein; regulated by acetylation, which modulates charge patch on its N-terminal tail; comprises about 20% of total H2A protein in Tetrahymena; found only in transcriptionally active nuclei.	
TTHERM_00476670	*RPL20*	Ribosomal Protein of the Large subunit #20; Homolog of Yeast RPL20, Human RPL18A; Ribosomal L18ae protein family protein.	2
TTHERM_01528510		Hypothetical protein.	2
TTHERM_00487110	*PRE10*	20S proteasome core alpha subunit 7.	
TTHERM_00375130		Carboxyvinyl-carboxyphosphonate phosphorylmutase-related.	
TTHERM_00895810		Hypothetical protein.	3
TTHERM_01029940		Hypothetical protein.	4
TTHERM_00075670	*RPL43*	Ribosomal Protein of the Large subunit #43; Homolog of Yeast RPL43, Human RPL37A.	
TTHERM_01014740		Hypothetical protein.	
TTHERM_01113100		40S ribosomal protein RACK1 (Receptor for Activated C Kinase 1); homolog of yeast ASC1.	2
TTHERM_00068170	*TCB2*	*Tetrahymena* Calcium-Binding protein; 25 kD Ca(2+)-binding protein containing four EF-hand loops; localized to cell cortex except in and around basal bodies, and around both the migratory and the stationary gametic pronuclei during pronuclear exchange.	
TTHERM_00445980		Metallopeptidase family M24 containing protein.	3

TTHERM_00926980		Acetyl-CoA acyltransferases family protein.	
TTHERM_00068120		Hypothetical protein, acyltransferase domain.	
TTHERM_00028740	*RPL12*	Ribosomal Protein of the Large subunit #12; Homolog of Yeast RPL12, Human RPL12, Bacterial RPL11; Ribosomal protein L11; RNA binding domain containing protein.	
TTHERM_00467660	*RPS24*	Ribosomal Protein of the Small subunit 24; Homolog of yeast RPS24, human RPS24.	
TTHERM_01100380		Hypothetical protein.	
TTHERM_00068140		Succinyl-CoA synthetase, alpha subunit family protein.	
TTHERM_00522720	*UBC9*	SUMO-conjugating enzyme involved in the Smt3p conjugation pathway; a member of the ubiquitin-conjugating enzyme E2 family.	2
TTHERM_00420030		Phosphoenolpyruvate carboxykinase.	
TTHERM_00216150		Hypothetical protein, Telomere stability and silencing; telomere stability C-terminal domain; Sde2 domain.	
TTHERM_00333210	*RPL4*	Ribosomal Protein of the large subunit; ribosomal protein L4/L1 family protein	3
TTHERM_01108700	*PGI1*	Glycolytic enzyme phosphoglucose isomerase, catalyzes the interconversion of glucose-6-phosphate and fructose-6-phosphate.	
TTHERM_00571860	*ATP1*	ATP synthase F1; alpha subunit family protein.	
TTHERM_00535530		Hypothetical protein.	2
TTHERM_00522810		Hypothetical protein.	2
TTHERM_00047080		Succinate dehydrogenase; flavoprotein subunit containing protein.	2
TTHERM_00250870		Dual specificity phosphatase; catalytic domain containing protein.	
TTHERM_00011390		Hypothetical protein.	4
TTHERM_00294870		Hypothetical protein.	

TTHERM_01289110	*RPL15*	Ribosomal protein of the large subunit; Homolog of yeast RPL15 (/RPL15e).	
TTHERM_00053800		FAD binding domain containing protein.	
TTHERM_00128280	*DFB1*	Deep fiber bundle protein 1; localizes to the deep fiber bundle of the oral apparatus, forms filamentous structures, associates with basal bodies and longitudinal microtubules.	
TTHERM_00225860		C2 domain. This repeat is found in a wide variety of proteins and generally consists of the motif XYPPX where X can be any amino acid.	2
TTHERM_01109770	*RPS11*	Ribosomal Protein of the Small subunit 11; Homolog of yeast RPS11, human RPS, bacterial RPS17.	
TTHERM_00152140		Brix domain containing protein. Proteins from the Imp4/Brix superfamily appear to be involved in ribosomal RNA processing, which essential for the functioning of all cells.	
TTHERM_00780640		Hypothetical protein.	
TTHERM_00667010		HSA domain (domain is predicted to bind DNA and is often found associated with helicases).	
TTHERM_01146030	*CAT1*	Catalase; peroxisomal enzyme involved in the metabolism of hydrogen peroxide; enzyme activity induced by linolenic acid.	2
TTHERM_00047230		Myb/SANT-like DNA-binding domain.	
TTHERM_00047580		Uroporphyrinogen decarboxylase family protein.	
TTHERM_00393040		Hypothetical protein.	
TTHERM_00859260		EF-1 guanine nucleotide exchange domain containing protein. Eukaryotic elongation factor 1 beta central acidic region; EF-1 guanine nucleotide exchange domain.	
TTHERM_00938820	*EEF2*	Eukaryotic translation Elongation Factor; catalyzes ribosomal translocation during protein synthesis; mRNA is expressed during vegetative growth; high amino acid identity with EF-2 in other eukaryotes	

TTHERM_00736480	*RPL5*	Ribosomal L18p/L5e family protein	
TTHERM_00499390		Glycine cleavage system T protein. Sequence similarity to the HisK protein kinase family. The system is a series of enzymes that are triggered in response to high concentrations of the amino acid glycine.	
TTHERM_01094890		Hypothetical protein.	
TTHERM_00245120		Hypothetical protein.	2
TTHERM_00852720		Zinc finger, C2H2 type.	2
TTHERM_01043090	*RPS27*	Ribosomal Protein of the Small subunit 27; Homolog of yeast RPS27, human RPS27	
TTHERM_00691610		CAP-Gly domain containing protein. CAP domain containing proteins are involved in organizing microtubules. The CAP-Gly domain is a conserved glycine-rich variant of the CAP domain. CAP-Gly domain bind to C-terminal sequence motifs present in α-tubulin and in other microtubule-associated proteins.	2
TTHERM_01299660		SCP-2 sterol transfer family protein. Sterol carrier proteins (also known as nonspecific lipid transfer proteins) is a family of proteins that transfer steroids and probably also phospholipids and gangliosides between cellular membranes.	
TTHERM_00971960		Hypothetical protein.	2
TTHERM_00044990		Fumarylacetoacetase family protein; FAA_hydrolase domain.	
TTHERM_00486310		Hypothetical protein.	
TTHERM_00066950		Pyridine nucleotide-disulphide oxidoreductase family protein.	
TTHERM_01097920		Hypothetical protein.	2
TTHERM_01001220		Hypothetical protein.	
TTHERM_01129660	*RPL36*	Ribosomal Protein of the Large subunit 36; Homolog of Yeast RPL36, Human RPL36; Ribosomal protein L36e containing protein.	

TTHERM_00487090		Fructose-bisphosphate aldolase class-I family protein.	
TTHERM_00522370		Cation channel family protein; ion transport protein.	2
TTHERM_00149300	*RPS4*	Ribosomal Protein of the Small subunit 4; Homolog of yeast RPS4, human RPS4.	
TTHERM_00773350		Hypothetical protein.	
TTHERM_00762890	*RPS17*	Ribosomal Protein of the Small subunit 17; Homolog of yeast RPS17, human RPS17.	2
TTHERM_00006320	*TTNA*	Tetrin A; novel, insoluble cytoskeletal protein with molecular mass of around 85 kD; unique to cilia.	
TTHERM_01020870		Zinc finger, C2H2 type family protein.	
TTHERM_00157949		Hypothetical protein.	
TTHERM_00357080	*SOD1*	Superoxide dismutase.	
TTHERM_00011400		Myb-like DNA-binding domain containing protein.	2
TTHERM_00227270	*RPL22*	Ribosomal Protein of the Large subunit #22; Homolog of Yeast RPL22, Human RPL22.	
TTHERM_01053030		CCT motif family protein.	
TTHERM_00402120		Glutathione S-transferase, N-terminal domain containing protein.	
TTHERM_00415630		Hypothetical protein.	
TTHERM_00299840		Methylmalonate-semialdehyde dehydrogenase family protein.	
TTHERM_00151850		Ser/Thr protein phosphatase family protein, Calcineurin-like phosphoesterase superfamily domain.	
TTHERM_00444800		Hypothetical protein.	
TTHERM_00134940	*RPL25*	Ribosomal Protein of the Large subunit #25; Homolog of Yeast RPL25, Human RPL23A, Bacteria RPL23; Ribosomal protein L23 containing protein.	
TTHERM_00379050	*PRE2*	20S proteasome core beta subunit 5.	

TTHERM_00499420	*TGP1*	G-quartet DNA-binding protein TGP1.	
TTHERM_00568050		HMG (high mobility group) box family protein.	
TTHERM_00361490		Hypothetical protein.	
TTHERM_00145440		Alpha/beta hydrolase family protein.	
TTHERM_01026240		Phytanoyl-CoA dioxygenase (PhyH).	
TTHERM_00518610	*DIM1*	Mitosis protein DIM1.	
TTHERM_00535700		RING-variant domain, FHA domain protein, putative.	
TTHERM_00824010		Ubiquitin-conjugating enzyme, putative.	
TTHERM_00046870		Phosphoglucomutase/phosphomannomutase, C-terminal domain containing protein.	
TTHERM_01035740	*MYO2*	Myosin heavy chain; member of proposed myosin Class XX, distinct from previously defined myosin classes; contains predicted coiled-coil, myosin tail homology (MyTH4), and FERM domains.	
TTHERM_00400750		D-hydantoinase family protein; Amidohydrolase family; Carbon-nitrogen hydrolase.	
TTHERM_00678260		*Tetrahymena* homolog of the liver F-Antigen; Homolog of vertebrate liver F-antigen; constituent of the intracellular membrane network; associated with membranes of the Golgi apparatus and transport vesicles.	2
TTHERM_00429890		Nucleoplasmin protein.	
TTHERM_00565590		Hypothetical protein.	

Table 8. Smt3p substrates identified from both native and denaturing conditions of lysis.

TGD ID	Gene Name
TTHERM_00339620	*RPL40*
TTHERM_00346560	
TTHERM_00105110	*HSP70*
TTHERM_01014750	
TTHERM_00216150	
TTHERM_00196370	*HSP60*
TTHERM_00852720	
TTHERM_01014740	
TTHERM_00499420	*TGP1*
TTHERM_00859260	
TTHERM_00140850	
TTHERM_01146030	*CAT1*
TTHERM_01029940	
TTHERM_00068140	
TTHERM_00415630	
TTHERM_00558440	*SSA5*
TTHERM_00429890	
TTHERM_00938820	*EEF2*
TTHERM_00245410	
TTHERM_00499390	

VITA

VITA

Amjad Nasir

2804 Horizon Dr Apt 3
West Lafayette, Indiana
(765) 414 4346
amjad.m.nasir@gmail.com
www.linkedin.com/in/amjadmnasir/

EDUCATION

Aug 2008 – May 2015	Ph.D.	Biochemistry	Purdue University
Jan 2004 – Jul 2007	B.S.	Biological Science	Lee University

PROFESSIONAL EXPERIENCE

Aug 2008 – May 2015 Research Assistant Purdue University

Completed PhD course and led research and development efforts for characterization of the post-translational protein modification SUMOylation in the ciliate *Tetrahymena*.

• Develop dual-affinity purification model in *Tetrahymena* for use in proteomics-based protein identification • Develop and validate novel protocols for use in ciliates • Qualitative and quantitative analysis of experimental data • Present research findings at regional and national scientific meetings and conferences • Set up Laboratory Information Management System for cloud-based storage of protocols and research data • Supervise undergraduate projects, develop experiments and analyze results, identify problems and recommend improvements • Ensure departmental safety standards are met

Jan 2010 – Dec 2014 Teaching Assistant Purdue University

Responsible for performing teaching and teaching-related duties to assist faculty members and professors.

• Develop plan of study for undergraduate laboratory courses • Teach undergraduate laboratory courses • Develop new laboratory experiments for undergraduate course • Grade quizzes, examinations and other course materials • Meet with students during office hours and provide individual tutoring to students

MEMBERSHIPS IN PROFESSIONAL SOCIETIES

American Association for the Advancement of Science (AAAS)
American Society for Biochemistry and Molecular Biology (ASBMB)
Member of Association of Southeastern Biologists (ASB)

PROFESSIONAL ACTIVITIES

2013 Purdue Graduate Student Government Senator

HONORS

Jan 2013 Michael Beach Travel Award, Purdue University
2004-2007 Presidential Scholarship

PRESENTATIONS

Jan 2015 Thesis Defense
Mar 2014 6th Thesis Committee Meeting
Nov 2013 Graduate Student/Post-doc Seminar Series
Feb 2013 5th Thesis Committee Meeting
Oct 2012 Graduate Student/Post-doc Seminar Series
Feb 2012 4th Thesis Committee Meeting
Apr 2011 Graduate Student/Post-doc Seminar Series
Apr 2011 Midwest Protozoology Society Meeting, Peoria, IL
Feb 2011 3rd Thesis Committee Meeting
Nov 2010 Preliminary Examination
Mar 2010 Graduate Student/Post-doc Seminar Series
Mar 2010 2nd Thesis Committee Meeting
Nov 2009 1st Thesis Committee Meeting
Feb 2009 Rotation Presentation (Liu Lab)
Dec 2008 Rotation Presentation (Tao Lab)
Sep 2008 Rotation Presentation (Forney Lab)
Apr 2008 Southwestern Association of Naturalists (SWAN) 2008 Annual Meeting, Memphis, TN
Apr 2008 Annual Meeting of the Association of Southeastern Biologists (ASB) and Annual Meeting of Tri-Beta National Biological Honor Society, Greenville, SC
Apr 2008 TAS Eastern Collegiate Division annual meeting, Harrogate, TN
Apr 2007 Annual Meeting of the Association of Southeastern Biologists (ASB) and Annual Meeting of Tri-Beta National Biological Honor Society, Columbia, SC
Dec 2006 Undergraduate Departmental Seminar

POSTERS

Aug 2014	Annual Biochemistry Retreat
May 2014	Seventh International Conference SUMO, Ubiquitin, UBL Proteins: Implications for Human Diseases Shanghai, China
Aug 2013	Annual Biochemistry Retreat
Jul 2013	FASEB Ciliate Molecular Biology, Steamboat Springs, CO
Apr 2013	Midwest Protozoology Society Meeting, Peoria, IL
Aug 2012	Annual Biochemistry Retreat
Apr 2012	Midwest Protozoology Society Meeting, Peoria, IL
Sep 2011	Annual Biochemistry Retreat
Oct 2010	Annual Biochemistry Retreat
Apr 2010	Midwest Protozoology Society Meeting, Peoria, IL

PUBLICATIONS

2015	Nasir AM, Yang Q, Chalker DL, Forney JD. SUMOylation is Developmentally Regulated and Required for Cell Pairing during Conjugation in *Tetrahymena*. *Eukaryotic Cell*.
2008	Nasir, Amjad M.; Okenkpu, Florence A.; Hisey, John R.; Atairu, Popoh A.; Kennedy, Michael L.; Carver, Brian D. A novel use of lineages in the analysis of effective population sizes of a solitary carnivore. *Journal of the Tennessee Academy of Science*.
2008	Nasir, Amjad M., Okenkpu, Florence A.; Hisey, John R.; Atairu, Popoh A.; Kennedy, Michael L.; Carver, Brian D. Raccoon movements, mating tactics, and effective population size. *Southeastern Biology*.
2007	Atairu, Popoh, Amjad Nasir, John Hisey, Jacqueline Famuyide, Uchechi Nwosu-Abanum, and Michael Kennedy. Group-Structured Genetics in a Solitary Carnivore. *Journal of the Tennessee Academy of Science*.
2007	Nwosu-Abanum, Uchechi, Amjad Nasir, Popoh Atairu, and John Hisey (2007) Group structured genetics in a solitary carnivore. *Southeastern Biology*.

TEACHING EXPERIENCES

Spring 2010	General Biochemistry Lab
Fall 2010	General Biochemistry Lab
Spring 2011	Medical Biochemistry Course
Spring 2011	Medical Biochemistry Lab
Fall 2011	General Biochemistry Lab
Spring 2012	Medical Biochemistry Course
Spring 2012	Medical Biochemistry Lab
Fall 2012	General Biochemistry Lab
Spring 2013	Medical Biochemistry Course
Spring 2013	Medical Biochemistry Lab
Summer 2013	General Biochemistry Lab

Fall 2013	Medical Biochemistry Course
Fall 2013	Medical Biochemistry Lab
Spring 2014	Biochemistry Lab
Summer2014	Biochemistry Lab
Fall 2014	Medical Biochemistry Course
Fall 2014	Medical Biochemistry Lab
Fall 2006	General Chemistry
Spring 2005	Microbiology