Identifying the Substrate Specificities of SENP1 and SENP2 in Recognition of Sumoylated Thymine-DNA Glycosylase

by

Yeong Je Jeong

A thesis submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science (ScM)

Baltimore, Maryland
April, 2018

Abstract

SUMO is an essential post-translational protein modification regulated in part by the activity of a family of SUMO-specific proteases known as SENPs. Mammalian cells express six different SENPs with essential and non-redundant functions. The molecular mechanisms that determine the substrate specificities of individual SENPs and their unique functions, however, remain unknown. Thymine-DNA glycosylase (TDG) is an important enzyme that recognizes and repairs G/U and G/T mismatches in the genome during the initial stages of base excision repair (BER), and this role is critical in genome integrity and also DNA demethylation. TDG is sumoylated *in vivo* and we are interested in exploring the functional importance and regulation of its modification. We previously found that SENP1 preferentially deconjugates sumoylated TDG *in vivo*, compared with SENP2, and that specificity is determined by the SENP1 catalytic domain (cSENP1). Here, we have used *in vitro* studies to further explore this specificity by comparing the activities of the SENP1 and SENP2 catalytic domains using multiple substrates, including AMC, RanGAP1 and TDG. Because TDG contains a SUMO interaction motif (SIM) that affects its modification *in vivo*, we hypothesized that non-covalent, intramolecular SUMO-SIM interactions impede TDG deconjugation. We also hypothesized that the cSENP1 is more efficient at disrupting the TDG SIM-SUMO interaction, thus explaining specificity. To test these hypotheses, we measured deconjugation rates of cSENP1 and cSENP2 over time for SUMO-modified RanGAP1, wild type and SIM mutant TDG. Our result supported a role for SIM binding in impeding

deconjugation, and also revealed interesting and unique substrate specificities for both cSENP1 and cSENP2.

Thesis Readers

Dr. Michael J. Matunis and Dr. Floyd Bryant; Department of Biochemistry and Molecular Biology, Johns Hopkins Bloomberg School of Public Health

Acknowledgements

Table of Contents

Introduction

Covalent modifications of proteins allow fast and energetically inexpensive alterations in

protein functions and these modifications, including phosphorylation, acetylation,

ubiquitylation, and sumoylation, regulate most cellular processes (Johnson, 2004;

Flotho & Melchoir, 2013). Sumoylation, one of the recently discovered post-translational

modifications, involves the covalent attachment of a 15 kDa small ubiquitin-related

modifier, or SUMO, to a target substrate (Boggio et al., 2004). Through this process,

sumoylation regulates a diversity of essential cellular processes, such as nuclear-

cytosolic transport, transcription regulation, cell cycle progression, as well as protein

stability (Hay, 2005). The covalent linkage occurs between the lysine residue of a target

substrate and the C-terminal glycine of the SUMO protein, and it is mediated by an

enzymatic cascade of activating enzyme E1 (SAE1/UBA2), conjugating enzyme E2

(Ubc9), and E3 ligase (Bernier-villamor et al., 2002; Reverter & Lima, 2005; Melchoir et

al., 2003). There are different SUMO paralogs, including SUMO1, SUMO2, and

SUMO3. SUMO2 and SUMO3 have ~96% identity and are collectively termed

SUMO2/3, while SUMO1 is ~45% identical to SUMO2/3 (Saitoh & Hinchey, 2000).

SUMO modification influences a target protein not only through covalent modification

but also through non-covalent interactions with SUMO-interacting motif (SIM) contained

on a target or another protein, mediating protein-protein interactions (Geiss-friedlander

& Melcoir, 2007). SUMO modification is reversed by Sentrin Isopeptidases, also known

as SENPs, and this regulation produces a dynamic cycling between conjugation and

deconjugation of the substrates (Mukhopadhyay & Dasso, 2007; Nayak & Muller, 2014).

1

There are six mammalian SENP paralogs, including SENP1, SENP2, SENP3, SENP5, SENP6, and SENP7, and they have diversified N-termini while sharing highly conserved catalytic domains.

Because sumoylation impacts proteins in ways ranging from changes in localization, altered activity and often stability of the sumoylated protein (Geiss-friedlander & Melcoir, 2007; Cubenas-Potts et al., 2013), the dynamic cycle of conjugation and deconjugation must be precisely regulated. Therefore investigating the unique SENP specificities for sumoylated substrates is a critical step in further understanding the nature of sumoylation and its influences. This is particularly important because an imbalance in the SUMO system contributes to initiation and progression of cancer and tumorigenesis (Bawa-khalfe & Yeh, 2010; Ei & Vertegaal, 2015).

Human thymine-DNA glycosylase (TDG) is an important enzyme in DNA repair, DNA demethylation, and transcription activation (Bellacosa & Drohat, 2015; Kohli & Zhang, 2013; Sjolund et al., 2013), and it is sumoylated. TDG is best known for its role in initial stages of base excision repair (BER) where it specifically recognizes G/U and G/T mismatches in the genome, generated from spontaneous deamination of cytosine or 5-methylcytosine, respectively. TDG proceeds to hydrolyze the N-glycosidic bond between the sugar phosphate backbone of DNA and the mispaired base, thus creating an abasic (AP) site (Barret et al., 1998; Lari et al., 2002). Interestingly, product inhibition at AP sites by TDG is observed, resulting in prevention of apurinic/apyrimidinic endonuclease 1 (APE1) from creating a single strand break and therefore unable to move on to the next steps of BER. Based on crystallography studies, sumoylation has

been proposed to induce a conformational change in TDG that alleviates product inhibition (Hardeland et al., 2002, Ulrich, 2003) (Figure 1).

TDG contains a SUMO-conjugation site at amino acid residue K330, and SIMs at residues 133-137 and 308-311. Once TDG is sumoylated at K330, by either SUMO1 or SUMO2/3, the attached SUMO interacts non-covalently with the SIM of TDG at E310 (Baba et al., 2005 & 2006; Smet-Nocca et al., 2011, Figure 1B). This intramolecular interaction leads to a conformational change in the N-terminus of TDG that results in the protrusion of the alpha-helix, and this is proposed to disrupt TDG binding to the AP site by decreasing DNA-binding affinity (Smet-nocca et al., 2011). However, the role of TDG sumoylation in alleviating product inhibition in BER was studied *in vitro* recently, revealing that sumoylation may not be sufficient (Coey et al., 2014). This study also revealed that the presence of APE1 may be sufficient to alleviate product inhibition. In addition, previous studies from the Matunis lab investigated the effects of sumoylation and SIM-SUMO interactions on TDG activity *in vivo*, and found that TDG base excision repair activity did not require sumoylation (Mclaughlin et al., 2016). Interestingly during this study, a unique specificity of SENP1 for sumoylated TDG was observed compared with cSENP2. It was observed that over-expression of SENP1 completely deconjugated sumoylated TDG, however SENP2 could only deconjugate at lower efficiency. Chimeric SENP1 and SENP2 enzymes with swapped catalytic domains, confirmed that this specificity lies within the catalytic domain of SENP1 (cSENP1). The over-expression of a SENP2 chimera with the SENP1 catalytic domain completely deconjugated sumoylated TDG whereas a cSENP1 chimera could not.

In this study, we investigated the substrate specificities of cSENP1 and cSENP2 for TDG *in vitro* in intention to explain the phenomenon observed in our previous *in vivo* experiments. We hypothesized that non-covalent, intermolecular SIM-SUMO interactions impede deconjugation of TDG. We also hypothesized that the cSENP1 is more efficient at disrupting the TDG SIM-SUMO interaction, thus explaining the unique specificity of cSENP1 (Figure 2). We first confirmed our cSENP1 and cSENP2 proteins as properly functioning isopeptidases by conducting deconjugation assays with sumoylated AMC (7-Amino-4-methylcoumarin) molecule and compared results with previously published data (Kolli et al., 2011). Then the substrate, SUMO1-RanGAP1 419 (Ran GTPase-activating protein 1), was utilized in deconjugation assays. Due to its lack of a SIM, we could measure baseline or control deconjugation rates to compare with sumoylated TDG. Finally, deconjugation rates were measured using wild-type TDG (TDGWT) and also a TDG SIM-mutant variant (E310Q) to observe the impeding effect of SIM-SUMO1 interactions on deconjugation by cSENPs. We observed more efficient deconjugation rates in assays with the TDG SIM-mutant, confirming the impeding role of SIM-SUMO1 interactions on deconjugation. However, both cSENPs were influenced by SIM-SUMO1 interactions equally, indicating no specificity in disruption of the interactions by one enzyme over another. Additionally, we found unique preferences of cSENP1 for the TDG SIM-mutant and RanGAP1 over TDGWT and cSENP2 for RanGAP1 over TDG substrates, and we propose unique substrate specificities being determined by combined SUMO1 and substrate binding to cSENPs.

Materials & Methods

1. Preparation of SENP Catalytic Domains

DE3 Rosetta competent cells were transformed with plasmids pET28a-SENP1 (catalytic domain 419-644) and pET28a-SENP2 (catalytic domain 365-590). Both plasmids included 6 Histidine tags and kanamycin resistance gene. First, 1 ul of the plasmid (70 ng/ul) was added to 50 ul of DE3 cells and left on ice for 10 minutes. The cells were then heat-shocked at 42 °C in water bath then left on ice for 2 minutes. 450 ul of Super Optimal Broth (SOB) medium was added to the cells then the cells were incubated at 37 °C for 20 minutes for recovery. The whole 500 ul of the recovered cells were spread on chloramphenicol (34 ug/mL) & kanamycin plate for overnight incubation at 37 °C. One or two colonies were used to make starter culture then 1.5 L culture was incubated at 37 °C until optical density of 0.8 (600 nm GeneQuant 100, GE). The cells were induced with 0.5 mM IPTG for 4 hours at 30 °C then centrifuged at 4000 g-force for 20 minutes to pellets and flash-frozen. The frozen pellets were resuspended in Lysis Buffer A (20% sucrose, 20 mM Tris-HCl pH 8, 1 mM DTT, 350 mM NaCl, 20 mM Imidazole, 20 ug/mL lysozyme, 1:100000 Benzonase, 1 mM PMSF, and 0.1% NP-40) and left on ice for 30 minutes. The cells were lysed with French Press Machine twice between 10,000 and 15,000 psi, then centrifuged for 30 minutes at 60,000 g-force at 4 °C. The supernatant was collected to be incubated with 500 ul Ni-NTA Agarose, equilibrated with Wash Buffer A (20 mM Tris-HCl pH 8, 350 mM NaCl, 1 mM DTT, and 20 mM Imidazole), for 1 hour at 4 °C. The mixture was poured into the column then washed with 10 bed volumes of Wash Buffer A. To elute SENP catalytic domains, 20 mL of Elution Buffer A

(Wash Buffer A with 250 mM Imidazole) was used to collect in 4 fractions. Each fraction was diluted with 5 mL of Dialysis Buffer (Wash Buffer A with 0 mM Imidazole) to prevent precipitation. After checking protein concentration, the first fraction was dialyzed at 4 °C, consecutively three times, in Dialysis Buffer at 1:3000. For further purification, the proteins were purified via Gel Filtration Column (Superdex-200) with Size Exclusion Buffer A (20 mM Tris-HCL pH 8, 350 mM NaCl, and 1 mM DTT).

2. Preparation of SUMO1-RanGAP1 419

Co-transformation method is used to SUMOylate the substrates RanGAP1 419 and TDG *in vivo*. Approximately 350 ng of each of two plasmids, E1/E2/SUMO1 (includes chloramphenicol resistance gene) and pGEX6P1 GST-RanGAP1 419 (includes carbenicillin resistance gene), were added to BL21 competent cells and the previous transformation method was followed. The 500 ul of the recovered cells were spread on chloramphenicol (34 ug/mL) & carbenicillin (50 ug/mL) plate. About 20 colonies were used to inoculate starter LB culture of 100 mL then 4 L culture is incubated at 37 °C until optimal density of 0.8. The cells were then induced with 0.25 mM IPTG for 17 hours at 22 °C and were later centrifuged as previously described and flash-frozen. The pellets were resuspended in Lysis Buffer B (50 mM Tris-HCl pH 8, 150 mM NaCl, 0.5% Triton X, 1 mM DTT, protease inhibitor, lysozyme and benzonase) and set on ice for 30 minutes. The mixture was lysed and centrifuged as previously described, and the supernatant was collected to incubate with 500 ul of Glutathione beads, equilibrated with Lysis Buffer B without Triton-X and protease inhibitor, for one hour at 4 °C. The

complex was washed with Lysis Buffer B then with Cleavage Buffer (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM EDTA, and 1 mM DTT). The mixture was poured into the column and was washed with 10 bed volume of Wash Buffer B (50 mM Tris-HCl pH 8, 150 mM NaCl, 1 mM DTT). The 500 ul bead-protein complex was incubated with 200 units of PreScission Protease overnight at 4 °C and SUMO1-RanGAP1 without GST-tag was eluted in supernatant via centrifugation. For further purification, the protein was dialyzed in Dialysis Buffer B (Buffer B with 0 M NaCl) then was purified by anion exchange HPLC using Mono Q Column (GE) with IE-A Buffer (20 mM Tris-HCl pH 8, 0 M NaCl and 1 mM DTT) at flow rate of 1 mL/min with a gradient of 0-50% IE-B Buffer (IE-A Buffer with 1 M NaCl) over 2 hours.

3. Preparation of SUMO1-TDGWT and SUMO1- SIM-Mutant (E310Q) TDG

Co-transformation method described above was also used for *in vivo* SUMOylation of both TDGWT and SIM Mutant TDG. Site-directed mutagenesis was performed to generate E310Q construct (includes Histidine 6 tag and kanamycin resistant gene) from pET28-hTDG construct. Co-transformation protocol was identical as that of RanGAP1 and used E1/E2/SUMO1 plasmid with either of the WT and SIM-Mutant constructs. After incubation with Ni-NTA agarose beads, the agarose-protein mixture was poured over the column to be washed with Wash Buffer C (50mM Tris-HCl pH 8, 300mM NaCl, 20mM Imidazole and 10mM DTT). The proteins were then eluted with Elution Buffer B (Wash Buffer C with 150mM Imidazole) in 10mLs and the first two fractions are dialyzed

in IE-A Buffer. For further purification, the proteins were purified using the same anion exchange HPLC method as RanGAP1.

4. In Vitro Analysis of Isopeptidase Activity for SUMO1-AMC

100 ul of assay mixture with Assay Buffer I (25mM Tris-HCl pH8, 150mM NaCl, 2mM DTT and 0.1% Tween 20) containing 1.6 uM of SUMO1-AMC (Boston Biochem UL-758) and purified cSENPs (0.1 nM for cSENP1 & 1 nM for cSENP2) were added to the wells on 96-well-microplate (Greiner Bio-One 655207). The microplate was inserted into spectrofluorometer and the emission fluorescence at 460 nm is recorded (excitatory wavelength of 360 nm) every 20 seconds with a range of 50 at 37 °C. A standard curve was generated with fluorescence recordings from the wells containing only the AMC molecules at concentrations of 0 uM, 0.05 uM, 0.1 uM, 0.2 uM, and 0.5 uM, to convert fluorescence into molar concentrations for quantification purposes. The converted data was plotted into a graph that represents deconjugation rates of the substrate using Prism 6 software. The data was obtained in triplicate.

5. In Vitro Analysis of Isopeptidase Activity for Sumoylated RanGAP1, TDGWT, and SIM Mutant TDG

In 1.5 mL tubes (Denville), master mixes of 100 uls consisting Assay Buffer I, 1 uM of the substrate, and cSENPs (1 nM for cSENP1 & 20 nM for cSENP2) were incubated in 30C water bath. At different time points (0, 30, 60, 90, 120, 300, 600, and or 900 seconds), 10 uls of the mix was drawn and was added into new tubes containing 1X

Sample Buffer to quench the reaction. The tubes containing 13 uls of quenched

reactions were then boiled on heat block (VWR) at above 100 °C for 10 minutes, then

were left on ice for 3 minutes and quickly centrifuged. The reactions were loaded into

12.5% SDS-gels for electrophoresis then stained with Bio-Safe Coomassie G-250 Stain

(161-0786) for 1 hour. The gels were destained with MilliQ water overnight at room

temperature, then scanned with Odyssey Imaging Systems (LI-COR Biosciences) for

analysis and quantification using Image Studio. The quantified data was plotted into

graphs, using Prism 6 software, showing deconjugation rates. The data was obtained in

triplicate.

Results

1. cSENP1 deconjugates SUMO1-AMC more efficiently than cSENP2

SUMO1-AMC, or sumoylated 7-Amino-4-methylcoumarin, is a commercial sumoylated substrate for SENPs and AMC's characteristic as a fluorophore (excitation and emission at 360 nm and 460 nm, respectively) allows measurement of isopeptidase activity (Fig. 3A). Previous studies showed a kinetics data comparing isopeptidase activities of cSENP1 and cSENP2 for SUMO1-AMC, and demonstrated that cSENP1 is a far more efficient isopeptidase than cSENP2 in SUMO1-AMC deconjugation (Kolli et al., 2010). We utilized this finding to test our purified cSENPs as functional isopeptidases, by looking at relative activities of cSENP1 and cSENP2 in deconjugation of SUMO1-AMC.

We observed a consistent decrease in the percentage of SUMO1-AMC over time in both assays with cSENP1 and cSENP2 (Fig. 3B). The figure shows the the initial rates of deconjugation of substrates within the initial 15 percent of product formation, and the slopes are -0.0587 and -0.1281 for cSENP1 and cSENP2, respectively. Notably a 10 times higher concentration of cSENP2 was needed in the assays to allow for accurate activity comparisons. The normalized slope values, where the slope of cSENP2 is divided by 10, are -0.05857 and -0.01281, respectively. The normalized rate ratio of cSENP1 to cSENP2 was 4.57. Therefore, consistent with the previous study (Kolli et al., 2010), cSENP1 showed more efficient deconjugation of SUMO1-AMC than cSENP2.

2. cSENP1 deconjugates SUMO1-RanGAP1 more efficiently than cSENP2

From the SUMO1-AMC assays, we found that our cSENPs were properly active as isopeptidases, and used sumoylated RanGAP1 419 as the next substrate of study. Because RanGAP1 419 does not contain a SIM, and SUMO1 is expected to be freely accessible for isopeptidase recognition. Activity assays with RanGAP1 and SENPs therefore provided an adequate control model for comparison with TDG, which has SIM. SENP1 in general shows better isopeptidase activity for SUMO1-proteins than SENP2, therefore the assays with RanGAP 419 provide baseline of relative activity for our cSENPs (Mikolajczyk et al., 2007; Reverter & Lima, 2009; Kolli et al., 2010). We sought to record the deconjugation rates of cSENPs for SUMO1-RanGAP 419 for this purpose.

Previous *in vitro* studies showed that the amount of cSENP1 and cSENP2 required to deconjugate 50% of sumoylated RanGAP1 N (amino acids 418-588) in an hour was 3.4 nM and 10.8 nM, respectively (Reverter and Lima, 2009). Consistent with the previous study, our results showed faster deconjugation of SUMO1-RanGAP 419 by cSENP1 compared with cSENP2 (Fig. 4A & 4B). On lane 7 of Figure 4A (at 300 seconds), cSENP1 activity assays showed a complete deconjugation of the SUMO1-RanGAP 419 (30 kDa). However, cSENP2 activity assays revealed only partial deconjugation of SUMO1-RanGAP 419 at 300 seconds and also later time points (Fig. 4B). The intensities of the protein bands on coomassie gels were used to quantify the results, as shown on the lane 2 of Figure 4A, and the quantification was plotted into a

graph that shows the change in percentage of SUMO1-RanGAP 419 over time. The

quantification utilized the following equation: {(Modified)/(Modified + Unmodified)} x 100.

Figure 4C from our findings showed that cSENP1 is far more efficient than

cSENP2 in deconjugating SUMO1-RanGAP 419, requiring 91.75 seconds and 137.9

seconds, respectively, for 50% deconjugation. The regression equations were

generated with the data from the first 30 seconds of the incubation where 10-15% of

products were formed. The slopes representing deconjugation rates are -0.5449 and

-0.3624 for cSENP1 and cSENP2, respectively. Notably, 1 nM of cSENP1 and 20 nM

cSENP2 were used in the assays. These concentrations of the enzymes were based on

titrations that resulted in similar and comparable deconjugation of the substrate. In

addition, this ratio provided an optimal intensity of protein bands on coomassie gels for

quantification purposes. The normalized slopes are -0.5449 and -0.01812 for cSENP1

and cSENP2, respectively, and the ratio of the normalized slopes revealed a 30 fold

difference in activity.

3. SIM-SUMO1 interactions in TDG Impede deconjugation

We previously found that SENP1, but not SENP2, deconjugates sumoylated TDG

in vivo and that specificity is determined by the SENP1 catalytic domain (Mclaughlin et

al., 2016). From here, we hypothesized that this unique specificity is due to non-

covalent interactions between SUMO1 and the SIM of TDG that impedes SENP2

recognition and deconjugation. We further hypothesized that the catalytic domain of

SENP1 is more efficient at disrupting the TDG SIM-SUMO interaction, thus explaining the specificity.

To first test the impact of SIM-SUMO1 interactions on TDG deconjugation, we compared SUMO1-TDGWT to a SUMO1-SIM mutant variant of TDG by measuring deconjugation rates with our cSENPs. Our findings showed that both cSENP1 and cSENP2 prefer the SIM mutant variant of TDG over TDGWT (Figure 5E). The observed initial slopes of deconjugation by cSENP1 were -0.42 and -0.53 for TDGWT and SIM mutant, respectively. The slopes with cSENP2 were -0.049 and -0.062, and both cSENPs resulted in a higher rate of deconjugation for the SIM mutant. The Figures 5A ~ 5D show results from deconjugation assays and reveal protein bands for modified (80 kDa) and unmodified (65 kDa) versions of TDGWT and SIM mutant TDG, respectively, over time. From cSENP1 assays (Figures 5A & 5B), we observed faster rates of deconjugation in assays with the SIM mutant TDG, where almost all of the substrate was deconjugated after 600 seconds of incubation (last lanes of Figures 5A & 5B). In addition, cSENP2 assays (Fig. 5C & 5D) also showed a similar trend where SIM mutants were deconjugated faster than TDGWT. The last lanes of Figures 3 & 4 revealed that nearly 50% of the modified SIM mutant TDG was deconjugated whereas the TDGWT was significantly deconjugated less at the same time point. Overall, our results showed faster deconjugation of the SIM-mutant TDG over TDGWT with both cSENPs.

4. SIM-SUMO1 interactions in TDG equally affect cSENP1 and cSENP2

The assays with TDGWT and the SIM mutant TDG demonstrated the preferences of the cSENPs for SIM-mutant over TDGWT. However we originally hypothesized that this SIM-SUMO1 interaction would be disrupted better by cSENP1 than cSENP2.

To further test our hypothesis, we compared the initial deconjugation slopes of cSENP1 to cSENP2 for both substrates. The initial deconjugation slope for the TDG SIM mutant was divided by the initial deconjugation slope of TDGWT to generate the ratio. Notably, this SIM mutant-to-TDGWT ratio was 1.27 for both cSENP1 and cSENP2 (Figure 5E). Our findings therefore suggest that the SIM-SUMO1 interaction impedes deconjugation by both cSENPs with equal efficiency.

5. cSENP1, not cSENP2, deconjugates the TDG SIM-mutant similar to RanGAP1 419

The cSENP1 deconjugation assays for all three substrates (RanGAP1, TDGWT, and TDG SIM mutant) were combined to observe differences in deconjugation rates among the substrates. Our result in Figure 6A showed that the initial slopes of RanGAP1 and the TDG SIM mutant were -0.5449 and -0.5379, respectively. The ratio of the two slopes was 1.013, revealing that cSENP1 exhibits a very similar specificity for both substrates. On the other hand, a similar analysis of cSENP2 revealed that the initial slopes for RanGAP1 and the TDG SIM mutant where -0.3624 and -0.0645 (Figure

6B). The ratio of the two slopes for cSENP2 was 5.8, indicating much slower

deconjugation of the TDG SIM mutant compared to RanGAP1.

Discussion

A proposed *in vitro* model generated from crystallography structures of sumoylated TDG suggested that sumoylation functioned to alleviate product inhibition following base excision (Hardeland and Steinacher, 2002, Figure 1). However, previous studies from the Matunis lab investigated the effects of sumoylation and SIM-SUMO interactions on TDG activity *in vivo*, and found that TDG base excision repair activity did not require sumoylation (Mclaughlin et al., 2016). Interestingly, the unique specificity of cSENP1 for TDG, compared with cSENP2, was also observed in this study and we intended to define the variables dictating this phenomenon through the studies reported here.

In our experiments, we hypothesized that SIM-SUMO1 interactions in sumoylated TDG impedes deconjugation, and that cSENP1 more efficiently disrupts this interaction, explaining the *in vivo* phenomenon. Our *in vitro* experiments demonstrated that TDG SIM-SUMO1 interactions do impede deconjugation. Faster deconjugation rates were observed for SIM-mutant TDG compared to TDGWT with both cSENP1 and cSENP2, consistent with the first hypothesis.

However, the impact of SIM-SUMO1 interactions on TDG deconjugation rate was nearly identical for cSENP1 and cSENP2, therefore not supporting our second hypothesis that cSENP1 more efficiently disrupts the interaction compared to cSENP2. Since both cSENPs were affected equally, whether or not cSENPs play a catalytic role in disruption of the SIM-SUMO1 interaction remains to be further investigated in the future. There is a possibility that cSENP binding to TDG may induce a conformational

change leading to the disruption of the SIM-SUMO1 interaction and deconjugation. Therefore, crystallographic studies of enzyme-substrate structures will be helpful in elucidating the mechanism of TDG deconjugation.

Previously, X-ray crystallography studies (Reverter and Lima, 2004) revealed that, in addition to an extensive interface between the protease and the globular SUMO domain, cSENP2 also contacts the substrate lysine and SUMO consensus conjugation site during deconjugation. Site-directed mutagenesis in cSENP2 residues that are found to establish hydrophobic interactions with RanGAP1 residues resulted in less efficient deconjugation, suggesting the role of enzyme-substrate interactions in the substrate recognition and binding. Our results showed that cSENP1 has similar preferences for RanGAP1 and the SIM-mutant TDG, as it was able to deconjugate the SIM-mutant TDG at a similar rate when compared to RanGAP1 419. It may suggest that cSENP1 recognizes only SUMO1 for its isopeptidase activity, at least for these two substrates (Figure 7A). On the other hand, cSENP2 showed a unique preference for RanGAP1 relative to both TDGWT & the SIM mutant TDG. cSENP2 deconjugation rates for SIM-mutant TDG was not similar to RanGAP1 but more similar to TDGWT, suggesting that cSENP2 recognizes the two substrates differently, compared with cSENP1. Our Figure 7B demonstrates a proposed model of cSENP2 recognizing not only SUMO1 but also part of the substrate, RanGAP1 419, for isopeptidase activity. The model, consistent with the previous study, suggests that the interaction of cSENP2 with the substrate, in addition to with SUMO1, may enhance substrate binding and deconjugation.

Conclusion

In conclusion, we observed an impeding effect of SIM-SUMO1 interactions in deconjugation of sumoylated TDG. However, both cSENP1 and cSENP2 activities toward sumoylated TDG were equally enhanced by mutating the SIM, indicating no specificity in disrupting SIM-SUMO1 interactions. How, and if, SIM-SUMO1 interactions are disrupted prior to deconjugation must be studied to better understand its impeding effect. Additionally, cSENP1 and cSENP2 showed unique substrate specificities beyond SUMO1 recognition *in vitro*. We observed that cSENP1 showed similar preferences for RanGAP1 and SIM-mutant TDG, whereas cSENP2 showed a unique preference for RanGAP1 relative to both TDGWT & SIM-mutant TDG. Finally, we propose models of substrate recognition for cSENP1 and cSENP2 that may explain the difference in substrate preference. Further structural studies that explore the substrate recognition of cSENPs will help to further elucidate SUMO isopeptidase substrate specificities.

Figure Legends

Figure 1. Proposed model of TDG sumoylation alleviating product inhibition

During initial stages of BER, endogenous unmodified TDG (green) recognizes G/U or G/T mismatches (red) and hydrolyzes the mispaired base, producing an AP site. TDG remains tightly bound on the AP site, preventing the downstream enzyme APE1 (orange) to bind and proceed. This product inhibition is proposed to be alleviated by sumoylation (yellow) of TDG. TDG sumoylation induces conformational change that results in reduction of its DNA affinity and therefore dissociation from the AP site. The sumoylated TDG is deconjugated by SENPs and recycled. The image is adapted from Hardeland et al., 2002

Figure 2. Hypothesized model of cSENP1 interacting with sumoylated TDG

Figure in green represents sumoylated TDG with sumoylation site at K330 and SIM at E310. The attached SUMO1 (yellow) is non-covalently interacting with SIM of TDG. We hypothesize that SIM-SUMO1 interaction impedes deconjugation by cSENPs and that cSENP1 is more efficient at disrupting these interactions. Once SIM-SUMO1 interactions are disrupted, the SUMO1 protein is more accessible for deconjugation.

Figure 3. SUMO1-AMC deconjugation assays.

(A) SUMO1-AMC is a sumoylated commercial substrate which consists of isopeptide bond between SUMO1 and AMC. Once SUMO1 is deconjugated by cSENPs, the AMC molecule is capable of being excited at 360 nm wavelength and of emitting 460 nm wavelength, which can be detected by spectrofluorometer. This image is adapted from Kolli et al., 2010.

(B) Graph comparing the percentage of 1.6 uM SUMO1-AMC over time after incubation with 0.1 nM cSENP1 (red) vs with 1 nM cSENP2 (blue). SUMO1-AMC substrate was incubated with either cSENP1 or cSENP2 in 96 well plate, and excitation and emission wavelengths at 360 nM and 460 nM, respectively, are applied to measure the rate of deconjugation, every 20 seconds. The measured fluorescence are converted to protein concentrations first with the standard curve (not shown here) then the percentage out of the initial SUMO1-AMC concentration is calculated and plotted into the graph. The box shows the initial velocities of deconjugation within the first 10~15% product formation period. The data was obtained in triplicate.

Figure 4. SUMO1-RanGAP1 419 deconjugation assays.

(A) Coomassie gel scan of 1 uM SUMO1-RanGAP 419 (~30 kDa) incubated with 1 nM cSENP1 at time points 0, 30, 60, 90, 120, 300, 600, and 900 seconds. Over time, more unmodified RanGAP 419 (~15 kDa) are observed. Red boxes in lane 1 shows the method of quantification {unmodified/(unmodified+modified)}. This quantification is used to plot the data into graph shown in Figure 4C.

(B) Coomassie gel scan of 1 uM SUMO1-RanGAP 419 (~30 kDa) incubated with 20

nM cSENP2 at identical time points as Figure 4A. The descriptions are identical as

Figure 4A.

(C) Graph comparing the percentage of SUMO1-RanGAP 419 over time after

incubation with 1 nM cSENP1 (red) vs with 20 nM cSENP2 (blue). The protein bands on

gel scans are quantified as described in Figure 4A. The box shows the initial velocities

of deconjugation within the first 10~15% product formation period. The data was

obtained in triplicate.

Figure 5. SUMO1-TDGWT and SUMO1-SIM-mutant TDG deconjugation assays.

(A) Coomassie gel scan of 1 uM SUMO1-TDGWT (~80 kDa) incubated with 1 nM

cSENP1 at time points 0, 30, 60, 90, 120, 300, and 600 seconds. Over time, more

unmodified TDGWT (~65 kDa) are observed as well as an increase in free SUMO1 (~15

kDa). The quantification method is identical as RanGAP assays.

(B) Coomassie gel scan of 1 uM SUMO1-SIM-mutant TDG (~80 kDa) incubated with 1

nM cSENP1 at identical time points as Figure 5A. The descriptions are identical as

Figure 5A.

(C) Coomassie gel scan of 1 uM SUMO1-TDGWT incubated with 20 nM cSENP2 at

identical time points as Figure 5A.

(D) Coomassie gel scan of 1 uM SUMO1-SIM-mutant TDG incubated with 20 nM

cSENP2 at identical time points as Figure 5A.

(E) Graph comparing the percentage of SUMO1-TDGWT vs SUMO1-SIM-mutant TDG over time after incubation with cSENPs. The box shows the initial velocities of deconjugation within the first 10~15% product formation period. Comparison of these initial velocities show that in cSENP1 assays, SIM-mutant TDG (magenta) is deconjugated more efficiently than TDGWT (red). Similarly, in cSENP2 assays, SIM-mutant TDG (cyan) is deconjugated more efficiently than TDGWT (blue) than cSENP2. Notice the ratio of these initial velocities, comparing SIM-mutant to TDGWT, is 1.27 for both cSENP1 and cSENP2 assays.

Figure 6. Comparison of TDG and RanGAP assays.

(A) Graph comparing the percentages of 1 uM SUMO1-TDGWT (red), 1 uM SUMO1-SIM-mutant TDG (magenta), and 1 uM SUMO1- RanGAP1   419 (black) over time after incubation with 1 nM cSENP1. The box shows the initial velocities of deconjugation within the first 10~15% product formation period, and shows very similar velocities between SIM-mutant and RanGAP1   419.

(B) Graph comparing the percentages of 1 uM SUMO1-TDGWT (blue), 1 uM SUMO1-SIM-mutant TDG (cyan), and 1 uM SUMO1-RanGAP 419 (black) over time after incubation with 20 nM cSENP2. The box shows the initial velocities of deconjugation within the first 10~15% product formation period, and shows more similar velocities between TDG substrates than with SUMO1-RanGAP 419.

Figure 7. Proposed models of substrate recognition by cSENPs

(A) Model 1 shows cSENP1 recognizing the substrates, SUMO1-RanGAP 419 and

SUMO1-TDGWT without SIM-SUMO1 interactions, to cleave isopeptide bonds. The

model proposes that cSENP1 recognizes only the attached SUMO1 for isopeptidase

activity. RanGAP1   419 is modified at K524 and TDG is modified at K330. E310 on

TDG indicates SIM residue.

(B) Model 2 shows cSENP2 recognizing the substrate, SUMO1-RanGAP 419 to cleave

isopeptide bond. The model proposes that cSENP2 not only recognize the attached

SUMO1 but also interacts with the substate residues (red dotted lines) therefore

enhancing its isopeptidase activity.

Figure 1.

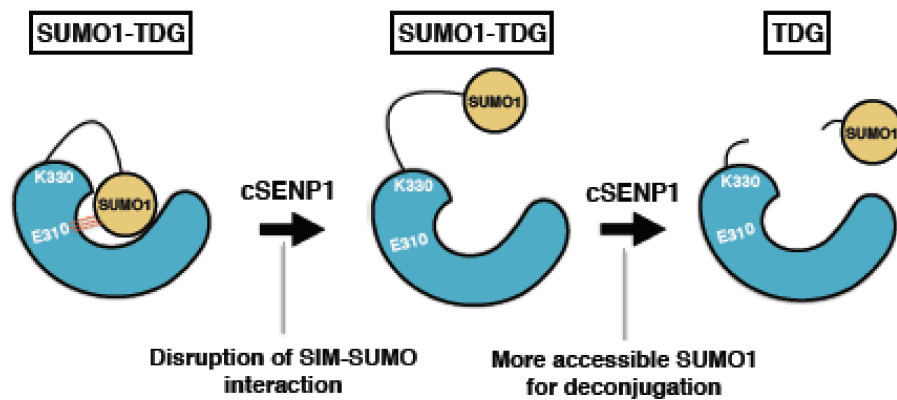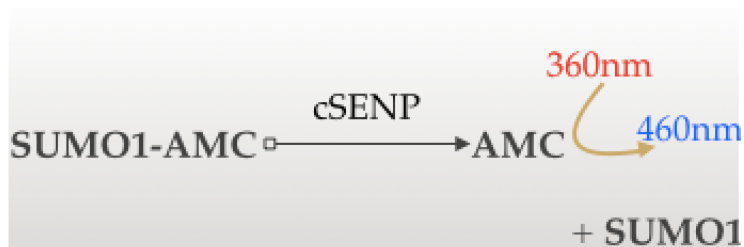Mismatch in DNA

TDG binds

Mismatch is removed &
AP site is produced

TDG
Sumoylation

Deconjugation
recycles TDG

TDG dissociates

continue
BER

Figure 2.

SUMO1-TDG

cSENP1

SUMO1-TDG

cSENP1

TDG

Disruption of SIM-SUMO
interaction

More accessible SUMO1
for deconjugation

Deconjugation of SUMO1-AMC (1.6 uM) by cSENPs

	cSENP1 (0.1nM)	cSENP2 (1nM)
Equation	Y = -0.05857*X + 98.50	Y = -0.1281*X + 102.2
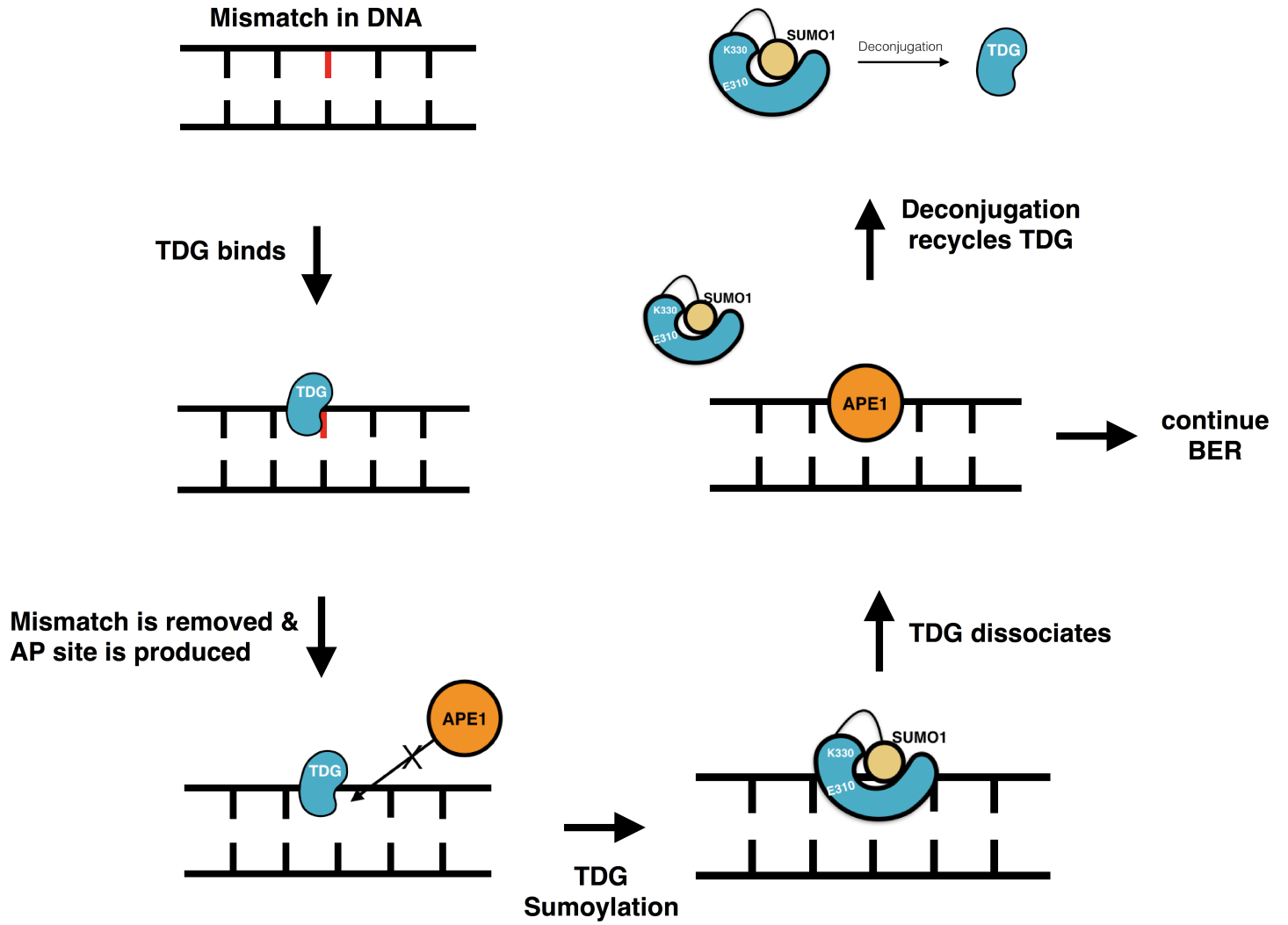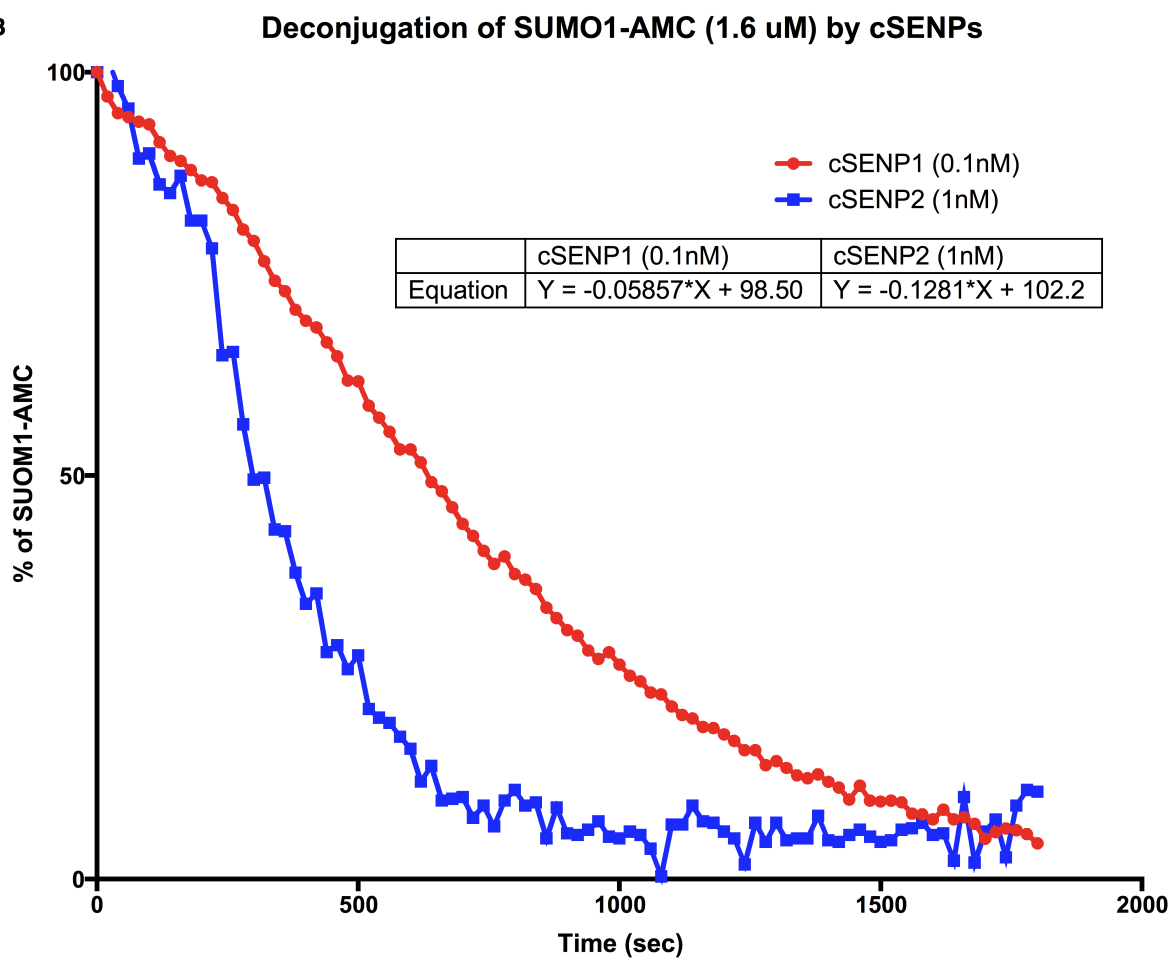
Figure 4A. SUMO1-RanGAP1 vs cSENP1

Incubation Time (sec)

SUMO1-RanGAP1
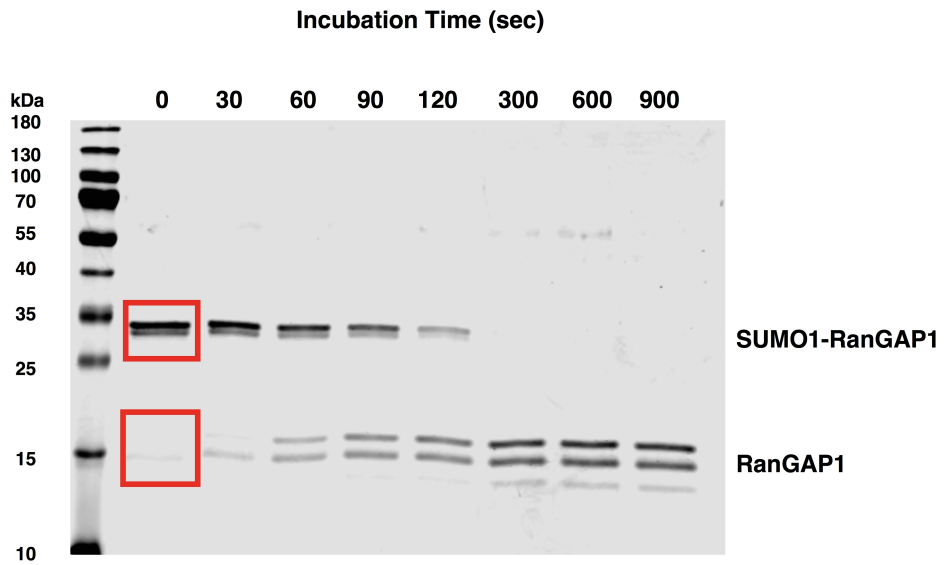
RanGAP1

Figure 4B. SUMO1-RanGAP1 vs cSENP2

Incubation Time (sec)

SUMO1-RanGAP1
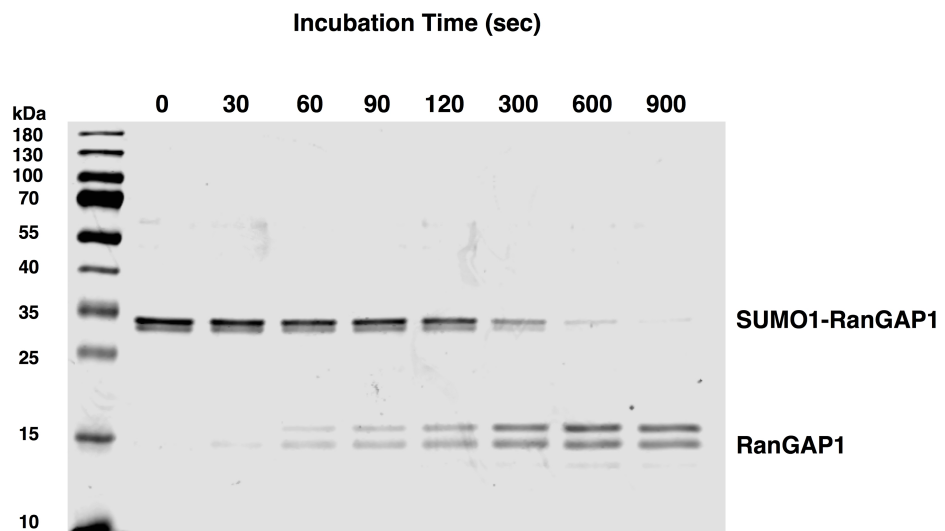
RanGAP1

Figure 4C.

cSENP1 vs cSENP2 in Deconjugation of SUMO1-RanGAP1

	cSENP1 (1nM)	cSENP2 (20nM)
Equation	Y = -0.5449*X + 100.0	Y = -0.3624*X + 100.0
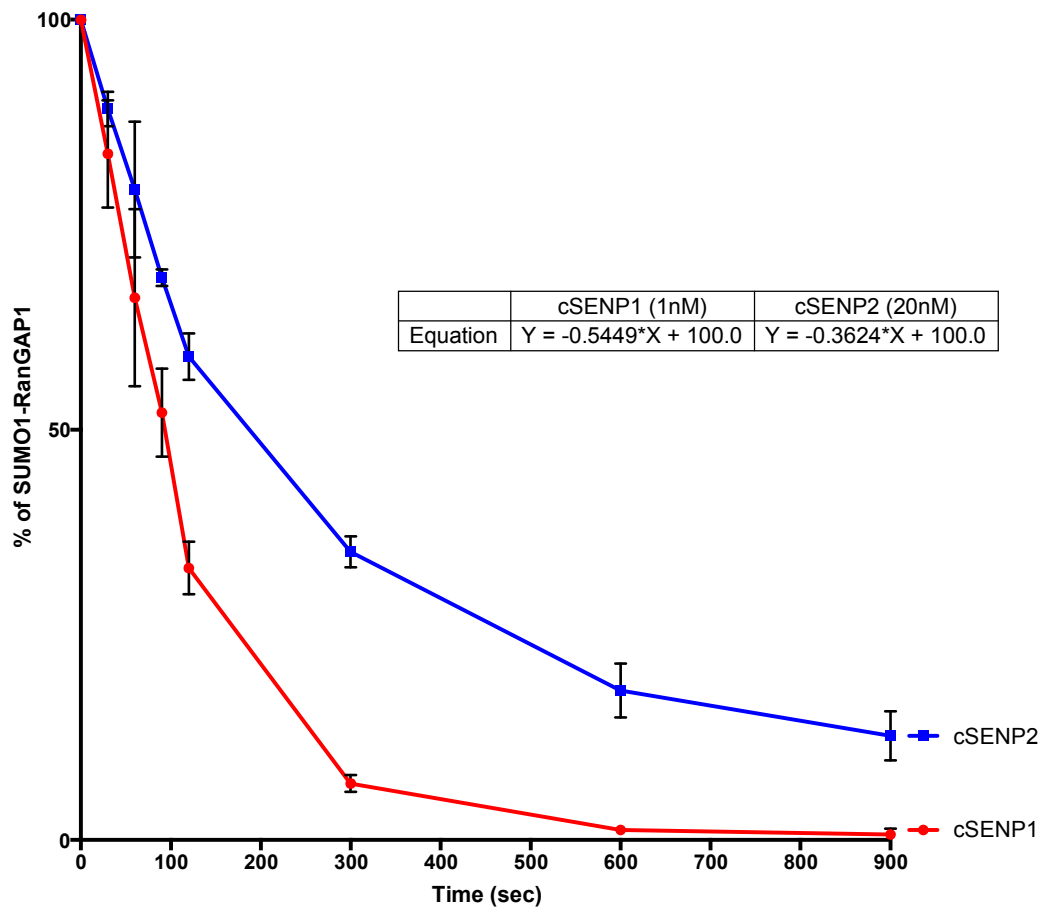
Figure 5A. SUMO1-TDGWT vs cSENP1

Incubation Time (sec)
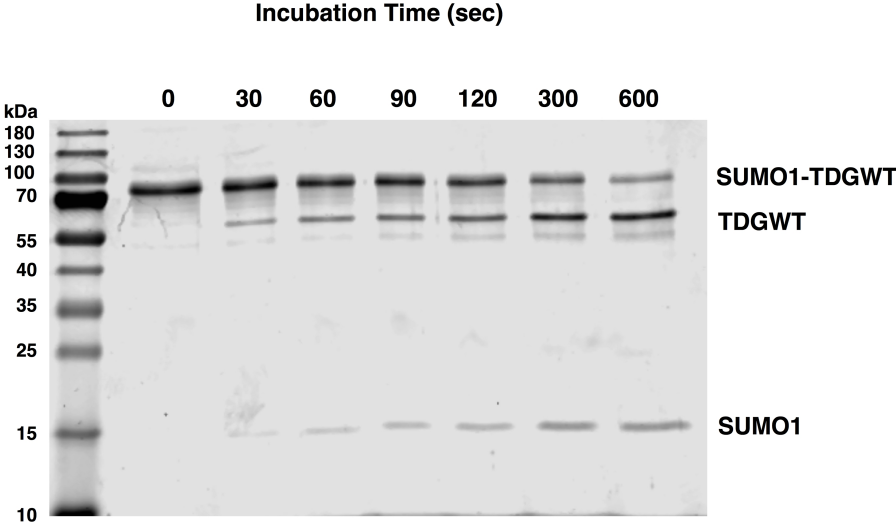
Figure 5B. SUMO1-E310Q-TDG vs cSENP1

Incubation Time (sec)
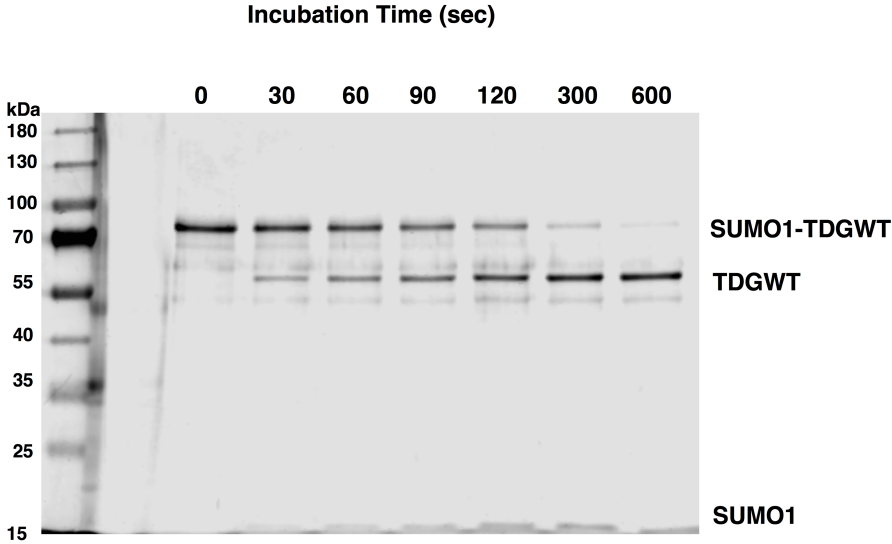
Figure 5C. SUMO1-TDGWT vs cSENP2

Incubation Time (sec)
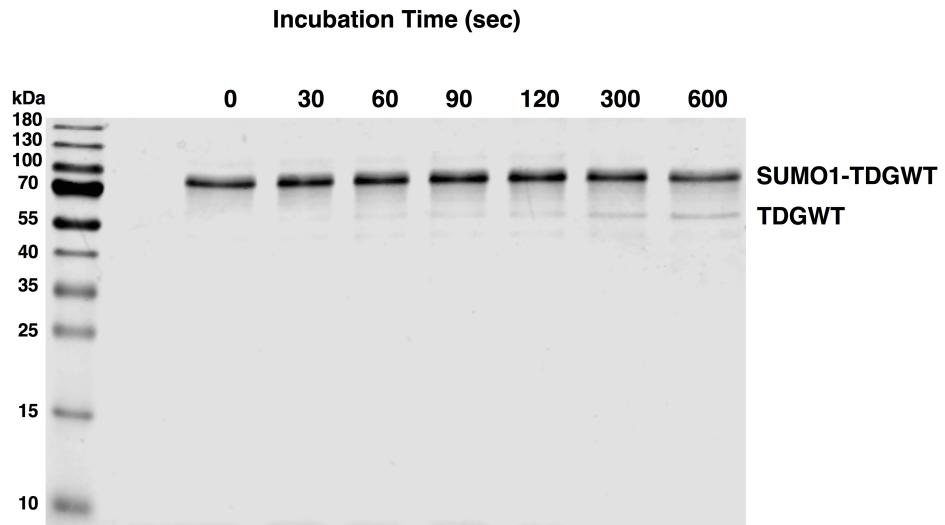
Figure 5D. SUMO1-E310Q-TDG vs cSENP2

Incubation Time (sec)
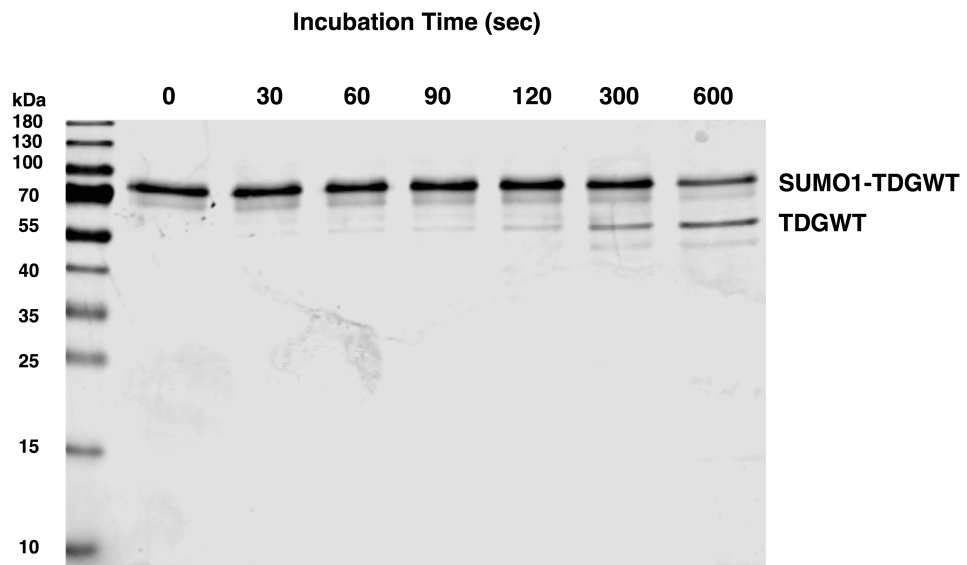
Figure 5E.

TDGWT vs SIM Mutant TDG Deconjugation

	Equation
SUMO1-TDGWT cSENP1	Y = -0.4209*X + 100.0
SUMO1-TDGWT cSENP2	Y = -0.04904*X + 100.8
SUMO1-SIM-Mut cSENP1	Y = -0.5379*X + 100.0
SUMO1-SIM-Mut cSENP2	Y = -0.06245*X + 99.56
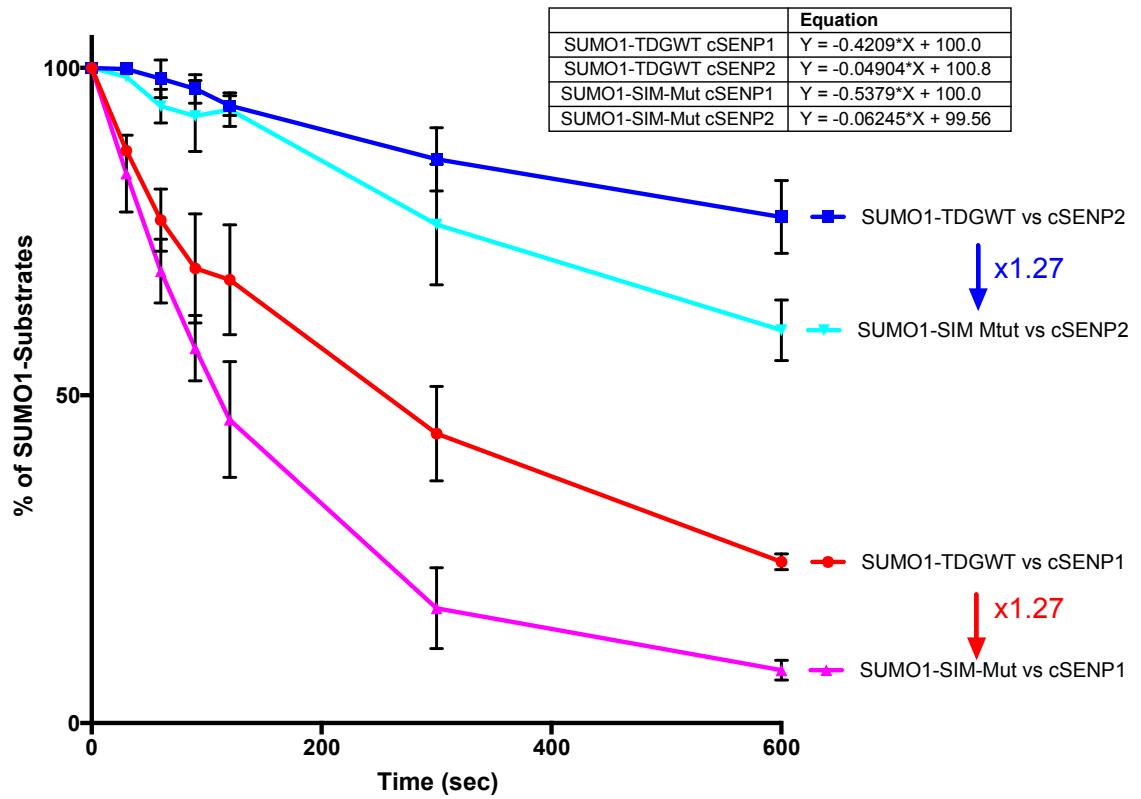
Figure 6A.   **Deconjugation Rates of Three Substrates by cSENP1**

	Equation
SUMO1-RanGAP1	Y = -0.5449*X + 100.0
SUMO1-TDGWT	Y = -0.4209*X + 100.0
SUMO1-SIM-Mut	Y = -0.5379*X + 100.0

SUMO1-TDGWT

SUMO1-SIM-Mut

SUMO1-RanGAP1
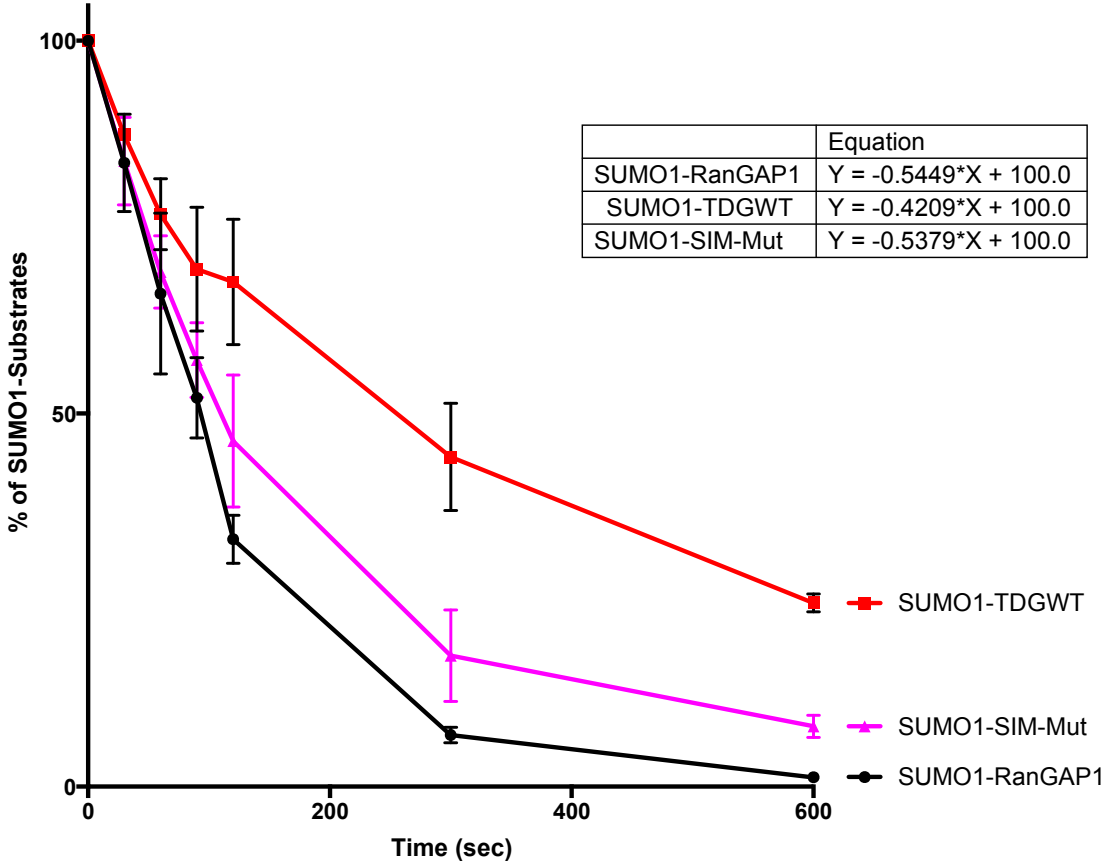
Figure 6B. Deconjugation Rates of Three Substrates by cSENP2

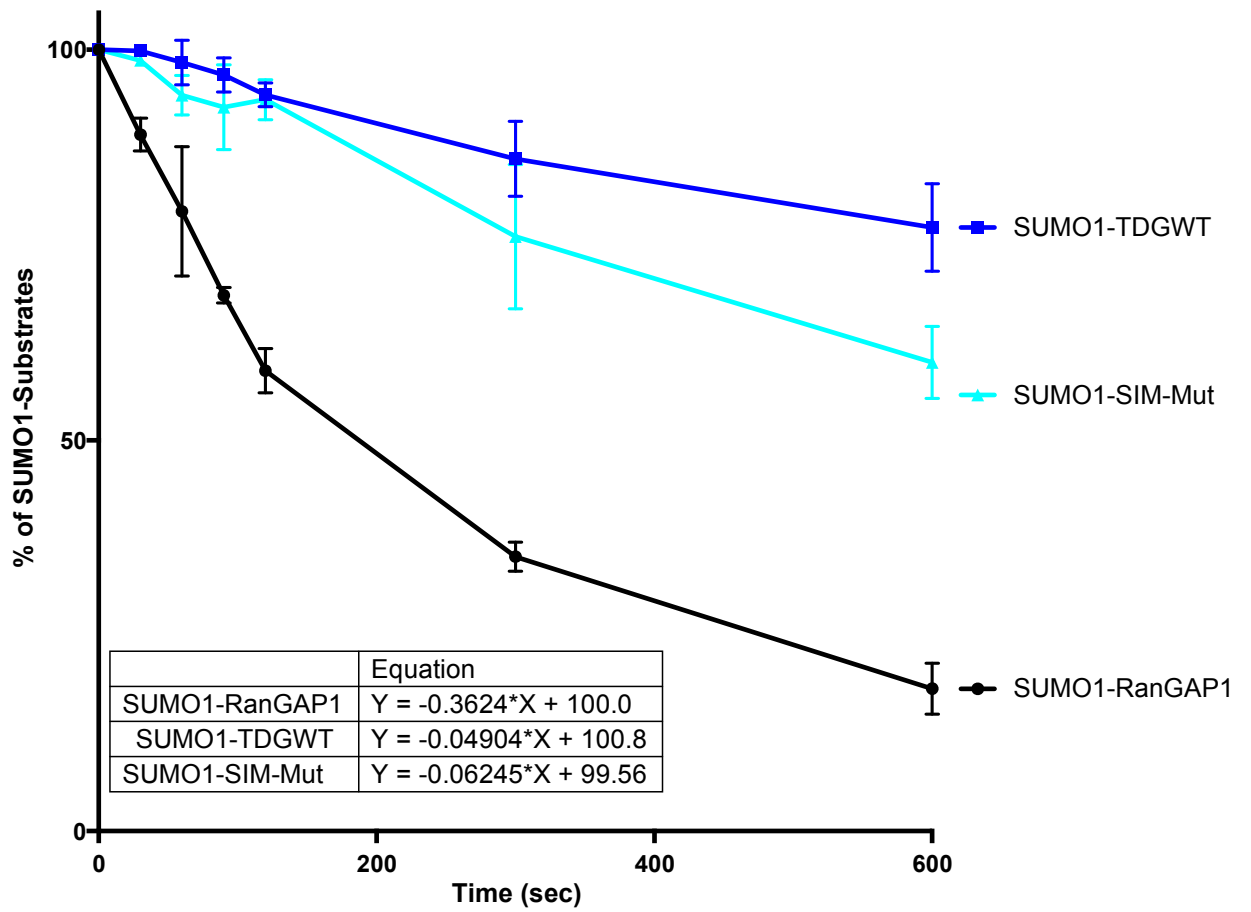|	Equation
SUMO1-RanGAP1	Y = -0.3624*X + 100.0
SUMO1-TDGWT	Y = -0.04904*X + 100.8
SUMO1-SIM-Mut	Y = -0.06245*X + 99.56

Figure 7A.
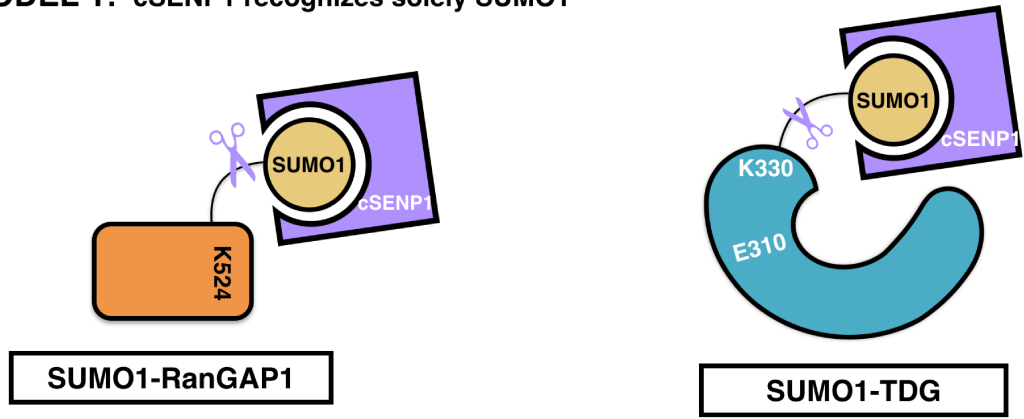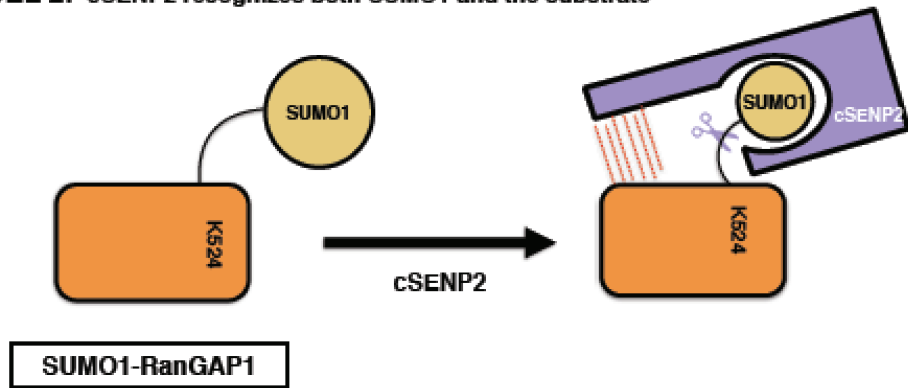

MODEL 1: cSENP1 recognizes solely SUMO1

Figure 7B.

MODEL 2: cSENP2 recognizes both SUMO1 and the substrate

33

References

Baba, D., Maita, N., Jee, J. G., Uchimura, Y., Saitoh, H., Sugasawa, K., . . . Shirakawa, M. (2005). Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. Nature, 435(7044), 979-982. doi: 10.1038/nature03634

Baba, D., Maita, N., Jee, J. G., Uchimura, Y., Saitoh, H., Sugasawa, K., . . . Shirakawa, M. (2006). Crystal structure of SUMO-3-modified thymine-DNA glycosylase. J Mol Biol, 359(1), 137-147. doi: 10.1016/j.jmb.2006.03.036

Barrett, T. E., Savva, R., Panayotou, G., Barlow, T., Brown, T., Jiricny, J., & Pearl, L. H. (1998). Crystal structure of a G:T/U mismatch-specific DNA glycosylase: mismatch recognition by complementary-strand interactions. Cell, 92(1), 117-129.

Bawa-khalfe, T., & Yeh, E. T. H. (2010). SUMO Losing Balance : SUMO Proteases Disrupt SUMO Homeostasis to Facilitate Cancer Development and Progression, 748–752. https://doi.org/10.1177/1947601910382555

Bellacosa, A., and Drohat, A. C. (2015) Role of base excision repair in maintaining the genetic and epigenetic integrity of CpG sites. DNARepair 32, 33–42.

Bernier-villamor, V., Sampson, D. A., Matunis, M. J., & Lima, C. D. (2002). Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1, 108, 345–356.

Boggio, R., Colombo, R., Hay, R. T., Draetta, G. F., & Chiocca, S. (2004). A Mechanism for Inhibiting the SUMO Pathway, 16, 549–561.

Cubenas-Potts, C., Goeres, J. D., & Matunis, M. J. (2013). SENP1 and SENP2 affect spatial and temporal control of sumoylation in mitosis. Mol Biol Cell, 24(22), 3483-3495. doi: 10.1091/mbc.E13-05-0230

Ei, K., & Vertegaal, A. C. O. (2015). SUMOylation-Mediated Regulation of Cell Cycle Progression and Cancer. Trends in Biochemical Sciences, 40(12), 779–793. https://doi.org/10.1016/j.tibs.2015.09.006

Flotho, A., & Melchior, F. (2013). Sumoylation: A Regulatory Protein Modification in Health and Disease. Annual Review of Biochemistry, 82(1), 357–385. https://doi.org/10.1146/annurev-biochem-061909-093311

Geiss-friedlander, R., & Melchior, F. (2007). Concepts in sumoylation: a decade on, 8(december), 947–956. https://doi.org/10.1038/nrm2293

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(6), 1456-1464. doi: 10.1093/emboj/21.6.1456

Hay, R. T. (2005). SUMO : A History of Modification Review, 18, 1–12. https://doi.org/10.1016/j.molcel.2005.03.012

Johnson, E. S. (2004). Rotein odification. https://doi.org/10.1146/annurev.biochem.73.011303.074118

J. Mikolajczyk, M. Drag, M. Bekes, J. T. Cao, Z., & Ronai, G. S. S. (2007). Small Ubiquitin-related Modifier (SUMO) -specific Proteases, 282(36), 26217–26224. https://doi.org/10.1074/jbc.M702444200

Kohli, R. M., and Zhang, Y. (2013) TET enzymes, TDG and the dynamics of DNA demethylation. Nature 502, 472–479

Kolli, N., Mikolajczyk, J., Drag, M., Mukhopadhyay, D., Moffatt, N., Dasso, M., … Wilkinson, K. D. (2010). Distribution and paralogue specificity of mammalian deSUMOylating enzymes, 344, 335–344. https://doi.org/10.1042/BJ20100504

Lari, S. U., Al-Khodairy, F., & Paterson, M. C. (2002). Substrate specificity and sequence preference of G:T mismatch repair: incision at G:T, O6-methylguanine:T, and G:U mispairs in DNA by human cell extracts. Biochemistry, 41(29), 9248-9255.

Matunis, M. J., Coutavas, E., & Blobel, G. (1996). A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. J Cell Biol, 135(6 Pt 1), 1457-1470.

Mclaughlin, D., Coey, C. T., Yang, W., Drohat, A. C., & Matunis, M. J. (2016). Characterizing Requirements for Small Ubiquitin-like Modifier (SUMO) Modification and Binding on Base Excision Repair Activity of Thymine-DNA Glycosylase in Vivo *, 291(17), 9014–9024. https://doi.org/10.1074/jbc.M115.706325

Melchior, F., Schergaut, M., & Pichler, A. (2003). SUMO : ligases , isopeptidases and nuclear pores, 28(11), 612–618. https://doi.org/10.1016/j.tibs.2003.09.002

Mukhopadhyay, D., & Dasso, M. (2007). Modification in reverse : the SUMO proteases, 32(6). https://doi.org/10.1016/j.tibs.2007.05.002

Nayak, A., & Müller, S. (2014). SUMO-specific proteases/isopeptidases: SENPs and beyond. Genome Biology, 15(7), 422. http://doi.org/10.1186/s13059-014-0422-2

Reverter, D., & Lima, C. D. (2004). A Basis for SUMO Protease Specificity Provided by Analysis of Human Senp2 and a Senp2-SUMO Complex, 12, 1519–1531. https://doi.org/10.1016/j.str.2004.05.023

Reverter, D., & Lima, C. D. (2005). Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. Nature, 435(7042), 687–692. http://doi.org/10.1038/nature03588

Reverter, D., & Lima, C. D. (2009). Preparation of SUMO proteases and kinetic analysis using endogenous substrates. Methods in Molecular Biology (Clifton, N.J.), 497, 225–239. http://doi.org/10.1007/978-1-59745-566-4_15

Saitoh, H., & Hinchey, J. (2000). Functional Heterogeneity of Small Ubiquitin-related Protein Modifiers SUMO-1 versus SUMO-2 / 3 *, 275(9), 6252–6258.

Sjolund, A. B., Senejani, A. G., and Sweasy, J. B. (2013) MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. Mutat. Res. 743–744, 12–25

Smet-Nocca, C., Wieruszeski, J. M., Leger, H., Eilebrecht, S., & Benecke, A. (2011). SUMO-1 regulates the conformational dynamics of thymine-DNA Glycosylase 48 regulatory domain and competes with its DNA binding activity. BMC Biochem, 12, 4. doi: 10.1186/1471-2091-12-4

Ulrich, H. D. (2003). SUMO Modification : Wrestling with Protein Conformation, (C), 257–259. https://doi.org/10.1016/j.cub.2005.03.027

Curriculum Vitae

101 N. Wolfe St #441
Baltimore, MD 21231
(410) 596-7597
yjeong11@jhu.edu

EDUCATION
2017-Present   *Sc.M. Candidate*, Biochemistry and Molecular Biology, **Johns Hopkins School of Public Health**, Baltimore, MD

2016-2017   *MHS*, Biochemistry and Molecular Biology, **Johns Hopkins School of Public Health**, Baltimore, MD

2004-2008   *B.S.*, Biological Sciences, **North Carolina State University**, Raleigh, NC

RESEARCH EXPERIENCE
2017-Present   Master's Student, Michael J. Matunis Lab, Department of Biochemistry and Molecular Biology, Johns Hopkins School of Public Health, Baltimore, MD
Principal Investigator: Michael J. Matunis, *Ph.D.*

HONORS AND DISTINCTIONS
2017-2018   Master's Tuition Scholarship, Department of Biochemistry and Molecular Biology, Johns Hopkins School of Public Health, Baltimore, MD

WORK EXPERIENCE
2012-2016   Real Estate Developer, Freelance, Seoul, South Korea

COMMUNITY SERVICE ACTIVITIES

2012-2016   Class Assistant, Miral School for Autism, Seoul, South Korea

2017-Present   Volunteer, Johns Hopkins Hospital, Baltimore, MD

2018-Present   Court Support Volunteer, Sanctuary Streets Baltimore Court Support, Baltimore, MD