Title
The regulation of global deSUMOylation by human SUMO- specific proteases

Permalink
https://escholarship.org/uc/item/9s35v5fh

Author
Békés, Miklós

Publication Date
2010-01-01

Peer reviewed|Thesis/dissertation
The Regulation of Global DeSUMOylation by Human SUMO-specific Proteases

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Molecular Pathology by Miklós Békés

Committee in charge:
Professor Guy S. Salvesen, Chair
Professor Christopher K. Glass, Co-Chair
Professor Alexandra C. Newton
Professor Susan S. Taylor
Professor Ze'ev Ronai
Professor Hudson H. Freeze
Professor Pieter C. Dorrestein

2010
The Dissertation of Miklós Békés is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

____________________________________

____________________________________

____________________________________

____________________________________

____________________________________

Co-Chair

Chair

University of California, San Diego

2010
DEDICATION

To my family

A Natáéknak, a Nagyiéknak,
Apunak,
Anyunak,
A Petinek
és az Erinnek
TABLE OF CONTENTS

Signature Page ........................................................................................................ iii
Dedication ................................................................................................................ iv
Table of Contents ...................................................................................................... v
List of Figures & Tables .......................................................................................... viii
Acknowledgements ................................................................................................. xi
Curriculum Vitae ..................................................................................................... xv
Abstract of the Dissertation ....................................................................................... xix

Chapter I
Introduction: SUMO-specific protease and their substrates,
the SUMO molecules .............................................................................................. 1

Abstract .................................................................................................................... 2
Introduction ............................................................................................................... 3

SUMO molecules in the context of ubiquitin-like modifiers ......................... 3
A decade of (de)SUMOylation history ............................................................. 4
The SUMOylation / deSUMOylation cycle .................................................... 6
The substrates of SENPs, SUMO molecules ............................................. 9
Deconjugating enzymes of the SUMOylation pathway .............................. 13

The family of SUMO-specific proteases ..................................................... 13
SENP specificity and catalytic mechanism ............................................... 13
SENP evolutionary origins ............................................................................. 16

The family of human SENPs ................................................................. 20
SENPas endopeptidases .................................................................................. 22
SUMO C-terminal extensions ................................................................. 26
SENPas isopeptidases ..................................................................................... 27
SENPas hiding in the nucleus? ................................................................. 30
SENPs in pathology ..................................................... 32
Open questions in SENP biology ................................. 35
Acknowledgements ..................................................... 37
References ............................................................... 38

Chapter II
Identifying (de)SUMOylation substrates by proteomics .... 51
Abstract ......................................................................... 52
Introduction ................................................................... 53
Results & Discussion .................................................... 56
DeSUMOMics – identifying lysines of deSUMOylation .... 56
DeSUMOMics of recombinant SUMOylated substrates .... 60
Identifying deSUMOylation substrates from cultured cells .... 66
SUMO-protoMAP identifies polySUMOylated substrates ..... 70
Materials & Methods ..................................................... 79
Acknowledgements ..................................................... 84
References ............................................................... 85

Chapter III
The mechanism and dynamics of SUMO chain deconjugation
by human SUMO-specific proteases ............................. 90
Abstract ......................................................................... 91
Introduction ................................................................... 92
Results .......................................................................... 95
The dynamics of SUMO chain deconjugation by
endogenous SENPs ....................................................... 95
Free and substrate-anchored SUMO chains are
substrates of SENPs ..................................................... 108
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear SUMO chains are cleaved in a stochastic manner</td>
<td>110</td>
</tr>
<tr>
<td>Cellular consequences of the lack of deSUMOylation</td>
<td>117</td>
</tr>
<tr>
<td>Discussion</td>
<td>123</td>
</tr>
<tr>
<td>The mechanism of SUMO chain deconjugation by SENPs</td>
<td>123</td>
</tr>
<tr>
<td>DeSUMOylation in vivo</td>
<td>127</td>
</tr>
<tr>
<td>The dynamics of SUMO chains</td>
<td>129</td>
</tr>
<tr>
<td>Materials &amp; Methods</td>
<td>131</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>138</td>
</tr>
<tr>
<td>References</td>
<td>139</td>
</tr>
</tbody>
</table>

**Chapter IV**

**Conclusions & Perspectives** 146

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outlook on the world of deSUMOylation</td>
<td>147</td>
</tr>
<tr>
<td>The future of SENP research</td>
<td>149</td>
</tr>
<tr>
<td>SENPs &amp; DUBs &amp; rock’n’roll</td>
<td>151</td>
</tr>
<tr>
<td>References</td>
<td>155</td>
</tr>
</tbody>
</table>
LIST OF FIGURES & TABLES

Chapter I

Figure 1.1 – The relationship between SUMO and ubiquitin …… 5
Figure 1.2 – The SUMOylation / deSUMOylation cycle ………… 7
Figure 1.3 – SENPs are dual function proteases ...................... 14
Figure 1.4 – The catalytic mechanism of SENPs ...................... 17
Figure 1.5 – Proteases of family C48: SENPs and ULPs .......... 21
Table 1.1 – DeSUMOylating enzymes in a nutshell ............... 28

Chapter II

Figure 2.1 – DeSUMOomics workflow ................................. 58
Figure 2.2 – PolySUMO2-RanGAP1 is modified by SUMO2 chains ........................................... 62
Figure 2.3 – Producing unanchored SUMO2 chains in E. coli .... 63
Figure 2.4 – Isopeptide linkages of SUMO2 chains .............. 64
Figure 2.5 – iTRAQ-ratios of a deSUMOomics experiment ........ 65
Figure 2.6 – DeSUMOomics of purified SUMO2-conjugates .... 68
Figure 2.7 – Isopeptidase specificity of ∆N-SENP1 versus ∆N-SENP6 ....................... 69
Figure 2.8 – SUMO-protoMAP workflow .......................... 72
Figure 2.9 – SUMOgraphs of proteins identified by SUMO-protoMAP – I .............................. 74
Figure 2.10 – SUMOgraphs of proteins identified by SUMO-protoMAP – II ............................ 75
Figure 2.11 – The utility of SUMO-protoMAP in visualizing ubiquitin chains ..................................... 78
Chapter III

Figure 3.1 – Enzymes of the SUMOylation pathway modeled with a P4 Gln to Pro (Q90P) SUMO2 mutant .......... 96
Figure 3.2 – SUMO2 Q90P is not deconjugated by endogenous SENPs ............................................. 98
Figure 3.3 – SUMO2 Q90P targets endogenous substrates ....... 99
Figure 3.4 – SUMO2 Q90P reveals the dynamics of SUMOylation in vivo .............................................. 100
Figure 3.5 – Purification of HMW-SUMO2-conjugates ............ 101
Figure 3.6 – HMW-SUMO2-conjugates in the Q90P SUMO2 fractions consist of polySUMO2 chains – I ........... 103
Figure 3.7 – HMW-SUMO2-conjugates in the Q90P SUMO2 fractions consist of polySUMO2 chains – II .......... 104
Figure 3.8 – Cleavage of HMW-SUMO2-conjugates by isopeptidases ...................................................... 105
Figure 3.9 – Cleavage of various length polySUMO2 chains ...... 106
Figure 3.10 – Full-length SENP1 also cleaves SUMO2 Q90P HMW-conjugates ........................................ 107
Figure 3.11 – Cleavage of recombinant SUMO chains by SENPs .. 109
Figure 3.12 – Linear tri-SUMO2 chains as a mechanistic model .... 111
Figure 3.13 – Cleavage of tri-SUMO2 constructs by ΔN-SENP1 .... 112
Figure 3.14 – Cleavage of tri-SUMO2 constructs by SENPs is stochastic ................................................... 113
Figure 3.15 – The loop within the catalytic domain of SENP6 is dispensable for activity and specificity ........... 114
Figure 3.16 – Full-length SENP1 cleaves tri-SUMO2 constructs ... 115
Figure 3.17 – *In vivo* effects of the lack of deSUMOylation in *S. pombe* ..................................................... 119

Figure 3.18 – Simultaneous knock-down of SUMO2 & SUMO3 in HEK293A cells .................................................. 120

Figure 3.19 – Simultaneous knock-down of SUMO2 & SUMO3 and concomitant reconstitution by wild-type & Q90P rescue plasmids ......................... 121

Figure 3.20 – In the absence of endogenous SUMOs, SUMO2 Q90P results in a slow-growth phenotype ... 122

Figure 3.21 – Modeling isopeptide-linked SUMO2 chains ............ 125

**Chapter IV**

Figure 4.1 – Schematic representation of the mechanism and the dynamics of SUMO chain deconjugation by SENPs ................. 148
ACKNOWLEDGEMENTS

It is difficult to begin my list of acknowledgements for I am quite humbled by the company of the great researchers I’ve had the pleasure to work with over these years here in San Diego. Guy Salvesen, however, stands out amongst the crowd, and not because he has allowed me to pursue the wildest projects, but because he has taught me how to think about science and has taught me how to be an independent scientist. His wit, insight and quick thinking on the spot are remarkable, making it a pleasure to work and compose manuscripts with him. I will forever be a protease-enthusiast no matter where life takes me, and for that I have to thank him.

The Salvesen lab has changed much since I started here one summer as an intern, but I will always remember it as the best place in – quite possibly, I must say – the world to be a graduate student. Coming here as a naïve undergraduate and thinking that graduate student years are going to be a piece of cake might not have turned out to be the whole truth, but every member of the Salvesen lab has made sure that the years pass fairly seamless. I am particularly grateful to Scott Snipas, to our unshakable lab manager who – am I sure he’s tired of hearing – is the most patient person to ever have walked the face of Clairemont Mesa, San Diego. Jean-Bernard Denault, an avid Chemical Brothers fan, has taken me under his wing at the beginning and has shown me the ropes of caspase-7 and of Anza-Borrego, both of which for I am eternally
grateful. Jowita Mikolajczyk has never tired of dealing with me and has got me hooked on SENPs, which eventually became my thesis project. The SENP team back then also consisted of Marcin Drag, a crazed organic chemist who became one of my best friends and with whom I thoroughly enjoy working with, talking about science with or simply attending conferences with. His support and friendship meant a lot these years. I am also grateful for the friendship and advice of Bram van Raam, who is frank and can always be counted on, no matter what. Cristina Pop, Fiona Scott and Mari Enoksson have always created a pleasurable atmosphere to work in, without their insider information I would know nothing of the dealings of the Burnham Institute. I also thank the support and the good times at happy hours of my fellow graduate students in the lab, Brendan Eckelman, John Timmer and Domenico Fasci, and everyone else in the greater Salvesen lab universe.

Graduate student life, however, is not confined to the walls of the lab. Therefore, there is a whole other list of people that I would like to thank. There are actually no words to describe how grateful I am for them, but I will try nonetheless. My parents, Békés Tamás and Békésné Hegyi Júlia, have been far away in kilometers throughout my graduate career, but in some ways have never been closer. None of them are scientists, yet the level of interest and encouragement they have shown is exceptional. They have raised me, and my brother, to be responsible adults who dare to make their own choices, and no matter what the consequences might be, to stand by them. I will love them
dearly forever, only hoping that one day I will be able to raise kids of my own with the love and support that they have raised me with. My brother, Péter, needs a special mention because, as the younger one, in my absence at home, he was the one who took care of mom and dad when I could not be there. Péter is a researcher as well, in the social sciences, and I wish him only the best in attaining his goals. Peto, imádlak!

During my graduate career I had the pleasure of becoming part of another family. They have taken me in with as much love as they have for their own daughter Erin, whom I married. Mike and Debbie Conn were there for me when my parents could not be and together with Erin’s brother Brad, have made California a home away from home for me. Finally, I would like to thank Erin, my awesome, awesome wife, who is a scientist herself – and a pretty good one too! Erin, you have made my life as easy-going and enjoyable as it can possibly be! Your reason and calm are my number one source of strength and I cannot wait to enjoy the rest of our lives together!

Apu, elmondani nem tudom, hogy mennyire szomorú vagyok, hogy nem érhetted meg, hogy befejezzem a PhD-t, vagy hogy a Peti befejezze a közgázt, vagy… Önzés tudom, de nagyon jó lenne ha itt lennél – pedig itt vagy, hát milyen hülyeség, nem!? – mert hiányzol nagyon, imádlak!

Chapter III was reproduced and modified with permission from co-authors from the manuscript: Békés, M., Prudden, J., Srikumar, T., Raught, B., Boddy, M.N., and Salvesen, G.S., The mechanism and dynamics of SUMO chain deconjugation by human SUMO-specific proteases, submitted for publication in 2010.
CURRICULUM VITAE

EDUCATION

2005 to 2010  Molecular Pathology PhD Program
University of California, San Diego, and Sanford-Burnham
Institute for Medical Research, La Jolla, CA

2002 to 2005  Biochemistry and Cell Biology B.Sc. Program
Jacobs (International) University Bremen, Germany

PUBLICATIONS

• Békés M, Prudden J, Srikumar T, Raught B, Boddy MN, Salvesen GS - The mechanism and dynamics of SUMO chain deconjugation by human SUMO-specific proteases, submitted 2010


• Békés M, Salvesen GS - The CULt of caspase-8 ubiquitination, Cell. 2009 May 15;137(4):604-6. (preview)


• Mikolajczyk J, Drag M, Békés M, Cao JT, Ronai Z, Salvesen GS - Small ubiquitin-related modifier (SUMO)-specific proteases: profiling the specificities and activities of human SENPs, J. Biol. Chem. 2007 Sep


RESEARCH EXPERIENCE

- Graduate student in Molecular Pathology PhD Program, University of California, San Diego: Research in the lab of Professor Guy S. Salvesen at the Sanford-Burnham Medical Research Institute, developing methods to profile the isopeptidase specificity of SUMO-specific proteases and determining the mechanistic details of SUMO chain deconjugation (September 2005 – present).

- Undergraduate researcher, Jacobs (International) University Bremen, Germany: Research in the labs of Professor Klaudia Brix and Professor Albert Jeltsch, examining cathepsin B secretion of rat thyroid cells and carrying out enzyme kinetic studies of a DNA-methyltransferase from E.coli, respectively (July 2003 – July 2005).

- Intern at the Burnham Institute: Summer internship in the lab of Professor Guy Salvesen, studying the activation mechanism of caspase-7 (June 2004 – August 2004).

- Intern at the Genome Damage and Stability Center at the University of Sussex, UK: Summer internship in the lab of Professor Alan Lehmann, studying the Smc5-Smc6 complexes in S. pombe (June 2003 – August 2003).

AWARDS & HONORS

- Half Moon Bay Award for the Best Talk at the 6th Annual Pacific Coast Protease Meeting, Warner Springs, 2009
• **International Proteolysis Society Travel Award** for the 5th meeting of the IPS, Surfer’s Paradise, Australia, 2009

• **International Proteolysis Society Travel Award** for the 4th meeting of the IPS, Patras, Greece, 2007

PRESENTATIONS


• **5th International Conference on SUMO, ubiquitin, UBL proteins. Implications for Human Diseases**, Houston, TX, February 2010. Poster presentation: *The Mechanism and Dynamics of SUMO Chain Deconjugation by Human SENPs.*


• **La Jolla Proteomics Conference**, La Jolla, CA, July 2008. Poster Presentation: *Profiling the Deconjugation Specificities of Human SENPs*

• **5th Annual Pacific Coast Protease Meeting**, Warner Springs, CA, April 2008. Oral presentation: *Determining the Regulation of Human SENPs*

• **4th International Conference on SUMO, ubiquitin, UBL proteins, Implications for Human Diseases**, Houston, TX, February 2008. Poster presentation: *The enigma of SUMO4.*
• 4th Meeting of the International Proteolysis Society, Patras, Greece, October, 2007. Poster Presentation: Deconjugation Profiles of Human DeSUMOylating Enzymes (SENPs)


• 10th Symposium on Proteases, Inhibitors and Biological Control, Portoroz, Slovenia, September 2006. Poster Presentation: Dissecting the Mechanism of Caspase-7 Dimerization

UNIVERSITY SERVICE

• Teacher Assistant in Histology, University of California, School of Medicine: teacher assistant under the guidance of Professor Andrew Mizisin for the Histology course required in the School of Medicine for second-year medical students and for first-year Molecular Pathology students (September 2006 – December 2006)
ABSTRACT OF THE DISSERTATION

The Regulation of Global DeSUMOylation by Human SUMO-specific Proteases

by

Miklós Békés

Doctor of Philosophy in Molecular Pathology, University of California, San Diego, 2010

Professor Guy S. Salvesen, Chair
Professor Christopher K. Glass, Co-Chair

SUMOylation – the covalent tagging of a target protein with a SUMO (small ubiquitin-related modifier) molecule – has recently emerged as an important post-translational modification in regulating a wide array of biological events. SUMOylation is a dynamics process, where SUMO-specific proteases (SENPs) process proSUMO molecules and also deconjugate them from substrates. My dissertation focuses on the deconjugating activity of SENPs and addresses the role they play in regulating the global SUMOylation “status quo” of eukaryotic cells.
SENPs are among the most specific proteases and only cleave SUMO molecules, yet they exhibit substantial substrate specificity among SUMO paralogs. To address the deconjugating specificities of human SENPs, I developed two different proteomics strategies to profile these proteases. One method has the potential to identify lysines of SUMOylation, besides the SUMOylation substrate itself, while the other can differentiate SUMOylated substrates from SUMO-interactors. To this end, I demonstrated the proof-of-principle of both methods and have identified polySUMOylated substrates.

Such polySUMOylation (i.e., modification of proteins by SUMO chains) has been an elusive phenomenon to study. In fact, so far, mainly mass spectrometry has shed light on the nature of SUMO chains through identifying SUMO-SUMO isopeptides. Thus, in the bulk of my dissertation, using a conjugatable, but not deconjugatable SUMO, I demonstrate the abundance of SUMO chains in vivo, make them visually appreciable and uncover the rapid dynamics of SUMOylation through their deconjugation. I propose that the balance of the dynamics is largely due to the action of SENPs in constantly cleaving SUMO from its targets, whether from a target protein or from another SUMO. This way, SENPs guard the SUMOylation status of the proteome, which by definition includes keeping in check modification of substrates by SUMO chains. I also establish the mechanism of SUMO chain deconjugation, which I show to be stochastic. Lastly, I reiterate the importance of deSUMOylation in vivo, by showing that the lack of deSUMOylation is adversary for the growth of Schizosaccharomyces pombe under replicative stress and for human cells.
Chapter I

Introduction:

SUMO-specific proteases and their substrates,
the SUMO molecules
ABSTRACT

Modification of proteins by SUMO is a dynamic and reversible process. The SUMO cycle begins when SUMO precursors are processed to remove short C-terminal extensions, thereby uncapping the C-terminal Gly-Gly motif that is essential for conjugation. SUMO ligases conjugate SUMO, via its C-terminal carboxylate, to the side-chain lysine of target proteins to generate an isopeptide linkage. Eventually, members of a protease family known as SENPs remove SUMO from SUMOylated proteins. SENPs, in common with other ubiquitin-like-specific proteases, are very limited proteases, and do not degrade their targets. They liberate the precursors from their C-terminal extensions (endopeptidase cleavage), and deconjugate SUMO from targets (isopeptidase cleavage). In these two ways they provide a pool of de novo and recycled SUMO for use by the ligation machinery. SENPs belong to a family of cysteine proteases, distinct from deubiquitinylating enzymes (DUBs). They show no sequence similarity to DUBs, although the fold has some similarities. To date seven SENPs have been identified in the human genome. They all contain a conserved C-terminal catalytic domain and have substantial N-terminal regions that show no similarity to any other protein domain and even resemble intrinsically unstructured proteins. This introductory chapter concentrates on SENP activity and their mechanism of deSUMOylation by describing current knowledge of how SENPs recognize their substrates, how they differ from other proteases, and how this correlates with their expected roles in biology and pathology.
INTRODUCTION

SUMO molecules in the context of ubiquitin-like covalent modifiers.
Before the deSUMOylating proteases (SENPs – sentrin/SUMO-specific proteases) are discussed their substrates, the SUMO (small ubiquitin-like modifier) molecules, need to be introduced in order to understand the importance of regulation carried out by the proteases. SUMO molecules are among a handful of small proteins that structurally resemble ubiquitin and also serves as covalent modifiers of their target proteins (Békés, 2010). The pathway of ligation and removal is analogous for UBLs (ubiquitin-like molecules), albeit carried out by specific conjugating and deconjugating enzymes. Among UBLs, besides ubiquitin itself, the SUMO molecules are the most studied and best understood modifiers and it is concerning the deconjugation of these molecules that my dissertation mainly tackles. Other UBLs less understood and quite possibly understudied are ISG15, an interferon-induced protein (Loeb and Haas, 1992), which is composed of two UBL domains and is involved in immunity, possible tagging newly synthesized viral proteins (Durfee et al., 2010); and FAT10, which is also has two UBL domains and is thought to cause proteasomal degradation of its conjugates (Hipp et al., 2005). On the other hand, Nedd8, which bears the most sequence identity to ubiquitin (Kamitani et al., 1997) and is involved in cullin E3 ligase regulation (Hori et al., 1999); together with Atg12, which is involved in regulating autophagy (Mizushima et al., 1998), have received considerably more attention than the previous two UBLs. Several
excellent reviews summarize the pathways, characteristics and proposed functions of many of the UBLs (Hochstrasser, 2009; Kerscher et al., 2006). However in this introductory chapter of my dissertation I will focus on SUMO molecules and the proteases that deconjugate them.

**A decade of (de)SUMOylation history.** Shortly after the discovery of sentrin, as a binding partner of Fas/APO-1 (Okura et al., 1996), the same 101 amino acid long polypeptide was identified as a covalent modifier of RanGAP1 (Mahajan et al., 1997). The protein was shown to be distantly related to ubiquitin in amino acid sequence but more so in structure and was termed SUMO, for small ubiquitin-like modifier (Figure 1.1). The discovery of SUMO as another post-translational modifier that is covalently conjugated to lysine side-chains of target proteins (in an analogous way to ubiquitinylation) paved the way for the discoveries of i) enzymes involved in attaching SUMO to target proteins, and ii) enzymes that remove SUMO from conjugated proteins. Within three years of the discovery of SUMO came the identification of the first protease shown to specifically deconjugate SUMO from target proteins (Li and Hochstrasser, 1999) and the study of deSUMOylating enzymes took off rapidly. This milestone study carried out by Li and Hochstrasser identified Ulp1 (Ubiquitin-like-specific protease 1) in *Saccharomyces cerevisiae* as a protease specifically cleaving yeast SUMO (Smt3), but not ubiquitin, from its conjugates. The study shed light on a whole new family of proteases and identified a crucial role of the first deSUMOylating enzyme in cell-cycle progression, underscoring the importance
Figure 1.1 – The relationship between SUMO and ubiquitin

A) Structure-based alignment of SUMO1 (in blue, PDB code: 1TGZ) and ubiquitin (in light grey, PDB: code 1UBQ), the image was generated using PyMol.

B) Sequence alignment of ubiquitin-like modifiers, produced using ClustalW (http://workbench.sdsc.edu). Fully conserved amino acids are indicated by a light blue star, partially conserved by a dark blue dot, similar amino acids by two green dots. The P4-P1 residues of the cleavage site, encompassing the conserved glycine residues utilized by ubiquitin-like proteases, are indicated in red. Internal lysines used for chain formation are in burgundy, with light blue inverted triangles indicating the seven lysines of ubiquitin, and a single dark triangle indicating the SUMOylation consensus site.
of regulated deSUMOylation as carried out by a deSUMOylating protease. It is imperative to mention at this point that deSUMOylating proteases have dual roles in the system, since they are responsible for both processing SUMO precursors to their conjugatable form and for deconjugating SUMO from target proteins. However, before this part of the chapter delves into discussing the mechanism, substrate-specificity and regulation of deSUMOylation by Sentrin/SUMO-specific proteases (commonly referred to as SENPs), I shall provide a brief introduction into the enzymatic machinery that covalently tags proteins with SUMO molecules.

**The SUMOylation / deSUMOylation cycle.** Since its discovery, SUMO has emerged as an important post-translational modification, regulating a plethora of biological processes, such as transcription (Gill, 2005), ribosome biogenesis (Panse et al., 2006), receptor function (Martin et al., 2007a), cell cycle control (Dasso, 2008) and DNA repair (Zhao and Blobel, 2005), to name but a few. Analogous to ubiquitinylation, SUMOylation is an energy-dependent process that tags target proteins with a SUMO moiety on the ε-amino-group of specific lysine side-chain residues. For the most recent advances in protein modification by SUMO with an emphasis on the SUMOylation side I refer to the excellent review by Geiss-Friedlander and Melchior describing a decade of SUMO research (Geiss-Friedlander and Melchior, 2007). As described briefly below (see Figure 1.2 for the scheme of the SUMOylation/deSUMOylation cycle), SUMO precursors need to be processed by an endopeptidase to remove their
Figure 1.2 – The SUMOylation / deSUMOylation cycle
Schematic representation of events in the SUMOylation/deSUMOylation cycle: (1) processing of the SUMO precursor by SENPs to its mature form exposing the C-terminal glycine, (2) activation in an ATP-dependent manner by the SUMO E1 activating enzyme, (3) transfer to the E2 SUMO-conjugating enzyme through a trans-esterification reaction, (4) ligation onto a lysine side-chain of a target protein by the E2 alone or through a SUMO E3 ligase, (5) deconjugation of SUMO from its substrate by SENPs.
C-terminal extensions, exposing the C-terminal glycine required for conjugation. The SUMOylation cycle is similar to ubiquitinylation in a sense that it requires the heterodimeric E1 SUMO-activating enzyme (Lois and Lima, 2005), and E2 conjugating enzyme (Song et al., 2005) and in some cases an E3 ligase (Johnson et al., 1997). There is a single active E1, the Aso1-Uba2 heterodimer (Lois and Lima, 2005), and a single active E2 enzyme, Ubc9 in both the human and yeast genomes (Johnson and Blobel, 1997). Ubc9 is capable of directly SUMOylating target proteins in vitro (Reverter and Lima, 2005), however there are also a handful of reported E3 ligases that presumably confer substrate specificity in vivo (Palvimo, 2007). Still the SUMOylation system appears to be far less complex than the attachment of ubiquitin to target proteins, which utilizes multiple E2s (Pichler et al., 2005) and hundreds of E3s (Petroski and Deshaies, 2005). SUMOylation also differs from ubiquitinylation in that there seems to be a consensus SUMOylation sequence, ψKXE (where ψ is a hydrophobic amino acid and X is any amino acid), within which the lysine is modified through an isopeptide-bond via the C-terminal glycine of SUMO (Hay, 2005), although some non-consensus SUMOylation sites have also been reported (Chung et al., 2004). Nonetheless, searching genomes for the SUMOylation consensus sequence is a powerful bioinformatics tool in trying to find SUMOylation substrates (Xu et al., 2008), although it could lead to many false positives as demonstrated by Wilkinson et al., who have shown that out of nearly forty predicted SUMOylation targets in neurons, less than half could actually be SUMOylated in vitro by SUMO1 (Wilkinson et al., 2008). On the other hand, a
recent proteomics study has shown that the majority of SUMO-conjugation sites identified by the authors do in fact conform to the consensus sequence (Matic et al., 2010).

Finally, the SUMO cycle is concluded with the removal of the SUMO moiety from the target lysine by an isopeptidase, recycling the SUMO molecule and making the lysine of the target protein available for a subsequent lysine post-translational modification. Again, the deSUMOylation machinery in humans is far less complex than the ubiquitinylation-deubiquitinylation system, which has upwards of 80 deubiquitinylating enzymes (DUBs), spanning several mechanistic classes (Reyes-Turcu et al., 2009; Wilkinson, 1997) as opposed to the six SUMO-specific proteases, SENPs, that begin and end the SUMO cycle (Mukhopadhyay and Dasso, 2007).

**The substrates of SENPs, SUMO molecules.** Lower eukaryotes posses a single copy of their SUMO homologue genes: the essential Smt3 (Johnson et al., 1997) and the non-essential Pmt3 (Tanaka et al., 1999) in *S. cerevisiae* and *S. pombe*, respectively (Johnson, 2004), whereas higher eukaryotes have multiple copies. For example, in humans four different SUMO genes (SUMO1-4) have been reported (Seeler and Dejean, 2003). However the capability of SUMO4 to conjugate target proteins has been questioned because of a unique proline residue near its cleavage/conjugation site (Owerbach et al., 2005). Furthermore it has recently been suggested to be a wrongfully assigned gene and may not be expressed at all (Bohren et al., 2007). SUMO molecules all
share a common fold with ubiquitin (Figure 1.1A) (Bayer et al., 1998), even though their sequence identity to ubiquitin is less than 20% (Figure 1.1B). Human SUMO2 and SUMO3 are more than 90% identical to each other, but are only 50% identical to SUMO1 (Geiss-Friedlander and Melchior, 2007). Importantly, they differ dramatically in their C-terminal tails providing specificity in SUMO maturation, the processing of these C-terminal extensions by SENPs (Hay, 2007; Mikolajczyk et al., 2007). Interestingly, although the precursors of SUMO2 and SUMO3 have radically distinct C-terminal extensions, once they have been conjugated they show little difference because their SUMO domains are almost identical. They become difficult to distinguish experimentally, and therefore the proteins they modify are routinely referred to as being SUMOylated by SUMO2/3. SUMO2 and SUMO3 are also capable of forming chains both in vitro and in vivo (Tatham et al., 2001), as demonstrated by an in vitro to in vivo proteomics strategy (Matic et al., 2008). Such polySUMOylation takes place through internal lysine side-chains that are found in the N-terminal regions of only SUMO2 and SUMO3, but are absent from SUMO1. Yeast Smt3 is also capable of forming chains in vivo, whose clearance is suggested to be the function of the second yeast SENP known as Ulp2 (Bylebyl et al., 2003). The functional consequences of polySUMOylation and the extent of mono- versus polySUMO modification of proteins have yet to be determined. Chapter III of my dissertation deals extensively with the abundance of SUMO chains in vivo and with the mechanism of their deconjugation by SENPs.
Although polySUMOylation is not associated with direct targeting of the substrate for proteasomal degradation like polyubiquitinylation, recent reports have described SUMO-dependent ubiquitin ligases (Prudden et al., 2007; Sun et al., 2007) that most likely recognize polySUMO chains through multiple SUMO-interacting motifs (SIMs) (Perry et al., 2008). Thus polySUMOylation could be an alternative signal for protein degradation by the ubiquitin/proteasome system (Lallemand-Breitenbach et al., 2008; Tatham et al., 2008). In this sense, deSUMOylation by SENPs could rescue the protein from recognition by ubiquitin ligases, thus would increase the half-life of the protein. On the other hand, for proteins where SUMOylation counteracts the proteasomal degradation signal, such as in the case of IκBα (Desterro et al., 1998), the action of SENPs would allow timely clearance of the substrate. Apart from the study of SENP1 knock-out mice (Cheng et al., 2007), the field has yet to see evidence for a specific function of a particular SENP involved in protein degradation. Although the multitude of SUMOylation substrates keeps increasing, there is currently very little information about the isopeptidase specificity of SENPs regulating spatial and temporal deSUMOylation of unique SUMOylated substrates.

Initially, SUMOylation was thought to be primarily involved in transcriptional repression, as many transcription factors have been shown to be negatively regulated by SUMO-modification (Gill, 2005), however today it is also widely accepted that SUMOylation alters the binding partners of the target proteins (Ling et al., 2004), resulting in various effects of protein SUMOylation. Even though many known SUMOylation substrates are nuclear proteins involved
in a wide array of functions such as chromosomal segregation, DNA-damage repair, maintenance of chromosome structure and transcriptional regulation, it has become clear from recent discoveries that SUMOylation is not unique to the nucleus and is also responsible for the modification of various cytosolic and plasma-membrane-bound proteins. This chapter focuses on general deSUMOylation mechanisms, however there are numerous excellent reviews covering the role of SUMOylation in the aforementioned processes (Gill, 2005; Huen and Chen, 2008; Martin et al., 2007b; Miura et al., 2007; Watts, 2007). Finally, with the identification of many extra-nuclear SUMOylated substrates, comes the question about the mechanism of deSUMOylation as carried out by – allegedly all nuclear – deSUMOylating enzymes, recycling SUMO and making the target protein available for a concomitant lysine post-translational modification.
DECONJUGATING ENZYMES OF THE SUMOYLATION PATHWAY

The family of SUMO-specific proteases. The dynamics of the SUMOylation cycle is achieved through the action of SENPs, an evolutionary conserved family of cysteine proteases that belong to peptidase clan CE, family C48 (Rawlings and Barrett, 1999). SENPs both activate SUMO precursors through their endopeptidase activity as well as remove SUMO from target proteins through their isopeptidase activity. These dual-function proteases are the main focus of this introductory chapter and they will be discussed in detail based on their mode of recognition of their substrates, their substrate specificity and their proposed regulation.

SENP specificity and catalytic mechanism. The substrate binding site in most proteases is composed of a groove that secures proper alignment of the substrate prior to hydrolysis. The binding site is divided into a number of sub-sites (Figure 1.3A), each securing a single amino acid residue of the substrate. In addition to interactions with specific side-chains, binding of the peptide backbone also plays an important role in catalysis. Together, these allow for specific interactions with the primary structure of the substrate (Schechter and Berger, 1967). SENPs, besides utilizing their active site for substrate recognition, also take advantage of a large exosite that is responsible for the binding of the SUMO moiety (Figure 1.3B). The same exosite is also responsible for the affinity of the substrate to the protease, as well as providing specificity.
Figure 1.3 – SENPs are dual function proteases
A) Schematic representation of the active site of proteases according to Schecter and Berger. P5-P4’ represent the amino acids of the substrate and S5-S4’ represent the specificity pockets of the protease. The hydrolyzed scissile bond is indicated in red. B) Top panel, the crystal structure of human SENP2 with proSUMO3 (PDB code: 2O11) in standard orientation showing the mode of SUMO binding to SENPs. The blow-up depicts how the (endopeptidase) substrate lies in the active site cleft. Lower panel, the crystal structure of human SENP2 with SUMO2-conjugated RanGAP1 (PDB code: 2IO3), showing how SENPs bind to SUMO-conjugated substrates. The blow-up again shows how the (isopeptidase) substrate lies in the active site pocket. Note how the mode of binding on the unprimed side (to the left of the scissile bond) is essentially identical for the endopeptidase and the isopeptidase substrate.
through hydrophobic interactions and allows SENPs to discriminate against UBLs other than SUMO molecules. On the other hand, the selectivity of SENPs towards the different SUMO precursors (paralog specificity) is dictated by specific interactions within the active site groove that allows SENPs to distinguish between the C-terminal extensions of SUMO molecules (Mikolajczyk et al., 2007; Reverter and Lima, 2006).

The enzymes that deconjugate UBLs are arguably the most specific of all proteases. SENPs and DUBs have high catalytic rates on their target substrates (SUMOylated and ubiquitinylated proteins, respectively) but have extremely low activity on other proteins, including short peptidyl substrates that encompass the cleavage site consensus for these proteases (Dang et al., 1998; Drag et al., 2008b). SENPs show no obvious sequence similarity to DUBs, but show a related protein fold. The DUBs that utilize a catalytic cysteine belong to Clan CA, and have an inverted orientation of the catalytic residues compared to Clan CE. This implies that Clans CA and CE share a common ancestor, and that the SENP fold might have undergone a domain-swap during evolution. The conserved catalytic residues of SENPs (family C48) are His514, Asp531, Gln574 with Cys580 being the active site cysteine used for hydrolysis (Ulp1 numbering according to Mossessova and Lima (Mossessova and Lima, 2000)). The catalytic Cys580 is coordinated by His514, which is itself stabilized by Asp531, to create the active cysteine thiol capable of attacking peptide bonds. The oxyanion hole, an essential component of most protease catalytic machineries used to in
destabilize the peptide-bond to allow hydrolysis, is formed by the main chain amide proton of Cys580 and the side chain amide proton of Gln574 (Figure 1.4).

**SENP evolutionary origins.** Presumably this exquisite specificity of SENPs represents a co-evolution of protease and substrate, and it comes as no surprise that distant SENP homologs are found throughout many primitive organisms. The protease-specific MEROPS database (merops.sanger.ac.uk) classifies SENPs as members of family C48 within Clan CE of the cysteine peptidases, based on sequence conservation within the catalytic domain (Barrett and Rawlings, 2001). The type example of family C48 is S. cerevisiae Ulp1, and members of this family are found in animals and plants, unicellular parasites (Ponder and Bogyo, 2007), bacteria (Hotson and Mudgett, 2004) and viruses (see Figure 1.5 for a schematic representation of the most studied members of family C48). Where information for these proteases is available, they all seem capable of targeting SUMO (or Nedd8, see below). More distant relatives, members of the parental Clan CE, may have unrelated substrates, the best known being adenain – the adenovirus processing protease. Adenain has the same fold and order of catalytic residues as SENPs, but its main activity seems to be to process viral coat protein precursors during viral assembly (Ruzindana-Umunyana et al., 2002), and as such its substrate specificity is substantially distinct from SENPs. Therefore it seems that the SENP protein fold (officially the Clan CE fold) can have varied catalytic specificity, dictated both by exosite and by active site groove interactions.
Figure 1.4 – The catalytic mechanism of SENPs
Schematic representation of the catalytic mechanism of SENPs showing how the active site cysteine attacks the destabilized peptide bond (red arrow). Catalytic side chains are in light grey. The substrate is in green sticks, and red and yellow sticks represent the conjugated lysine. The active site cysteine (Cys580) is deprotonated by a histidine residue (His514), which in turn is stabilized by an aspartate (Asp531), forming the catalytic triad of SENPs. Destabilization of the peptide bond is achieved by the main chain amide of the active site cysteine (Cys580) and the side chain amide of a glutamine residue (Gln574) forming the oxyanion hole (dashed olive lines).
As for the evolutionary conservation of deSUMOylating enzymes in family C48, the plant pathogen Xanthomonas campestris contains one gene, XopD, whose product has been shown to be a deSUMOylating enzyme. XopD, when injected into host plant cells, is capable of processing SUMOylated proteins in planta (Hotson et al., 2003). The crystal structure of XopD has also been solved providing evidence for its specificity towards SUMO proteins (Chosed et al., 2007). Together, with the crystal structures of the catalytic domains of Ulp1 from yeast (Mossessova and Lima, 2000), SENP1 (Shen et al., 2006b), SENP2 (Reverter and Lima, 2004) and SENP7 (Lima and Reverter, 2008) from humans, these structural studies provide insight into the mode of binding of SUMO to the protease. A database search using the catalytic domain of Ulp1 identified eight putative genes in Arabidopsis thaliana, out of which only two have been shown to have deSUMOylating activity against plant SUMO molecules, while one is Nedd8-specific (Colby et al., 2006). Baker’s yeast (S. cerevisiae) contains two deSUMOylating enzymes, Ulp1 and Ulp2, whereas fission yeast (S. pombe) only has one reported member of the family, Ulp1 or Nep1 (Taylor et al., 2002), and all are capable of deSUMOylation in vivo. Deletion of Ulp1 in S. cerevisiae is lethal and Ulp1 is thought to be the main Smt3 processing enzyme, required for the maturation of Smt3, making it available for conjugation (Li and Hochstrasser, 1999). On the other hand, Ulp2 is non-essential for viability, but is required for meiotic recovery and presumably serves more of a deconjugating/editing function being responsible for the remodeling of Smt3-modified foci in synaptonemal complexes (Schwartz et al., 2007). Ulp2 is also required for an
Smt3 chain editing function, however chain formation is dispensable for *S. cerevisiae* (Bylebyl et al., 2003).

SENPs are quite abundant in higher eukaryotes, ranging from single to multiple copies of deSUMOylating enzymes depending on the species. The MEROPS database contains the most up-to-date information about homologues of SENPs and since genes are often mis-assigned, it is always useful to check for “novel” proteases. For example, *C. elegans* currently has one Ulp1-like protease assigned in the C48 family, but has numerous other candidate genes whose deSUMOylating activity has yet to be demonstrated experimentally. *Drosophila melanogaster* also has only one deSUMOylating enzyme assigned, even though its activity *per se* has not been demonstrated. However, RNAi experiments in S2 cells suggest it has a function in deSUMOylating nuclear proteins (Smith et al., 2004). As a brief overview of vertebrates, *Brachydanio rerio* has six copies of SENPs, *Canis familiaris* has seven, while *Mus musculus* has more than fifteen members of the C48 family currently assigned. The apparent expansion of SENPs in the rodent lineage is unexplained, but may complicate assignment of functions of human SENPs based on mouse knockout studies – an obvious fact that investigators in the SUMO field tend to ignore.

For classification of SENPs, Mukhopadhyay et al. have shown that eukaryotic proteases cluster into different subgroups called Ulp1-like and Ulp2-like based on their sequence homology (Mukhopadhyay et al., 2006). This phylogenetic relationship presents a clear-cut view of evolutionary relationships between the two sub-groups of deSUMOylating enzymes, and in combination
with substrate-specificity data of human SENPs from our lab (Drag et al., 2008b) provides an excellent classification of the different activities of eukaryotic deSUMOylating enzymes. For further information about the abundance of SENPs throughout species please visit the MEROPS website at merops.sanger.ac.uk, which should always be a first stop when researching protease databases.

**The family of human SENPs.** Humans have seven SENPs in their genome, numbered SENP1-3 and SENP5-8 (SENP4 is thought to represent a pseudogene). SENP8, initially called NEDP1 (Mendoza et al., 2003) or DEN1 (Gan-Erdene et al., 2003), has been shown to be specific for Nedd8, another ubiquitin-like modifier more closely resembling ubiquitin, and does not cleave SUMO. The crystal structure of SENP8 has also been solved and provides evidence for its ability to distinguish between Nedd8 and ubiquitin based largely on a single amino acid (Shen et al., 2005). Similar to SUMO molecules, Nedd8 is produced as a precursor with an extended C-terminal tail that must be removed before conjugation, and SENP8 is capable of both removing the tail, and deconjugating Nedd8 from neddylated proteins (Wu et al., 2003).

DeSUMOylating SENPs 1, 2, 3, 5, 6 and 7 are comprised of two distinct regions, a C-terminal catalytic domain and an N-terminal regulatory region (Figure 1.5) which has no conserved fold and whose function is currently unknown, but is thought to be involved in directing the subcellular localization of SENPs and will be discussed in detail later in this chapter. The six SUMO-
Figure 1.5 – Proteases of family C48: SENPs and ULPs
Schematic representation of the most well characterized members of protease clan CE, family C48. The conserved catalytic cysteine is highlighted white within the conserved catalytic domain (dark green) which defines peptidase family C48. The conserved fold is called the Ulp-fold. N-terminal regions of SENPs (light red) are largely unstructured – based on secondary structure predictions – and show no homology to other known protein domains. Notice the insertions (purple) within the catalytic domains of SENP6 and SENP7. The subgroups in the family are determined based on phylogeny and SUMO paralog specificity.
specific human SENPs, which have an overall sequence identity of 20-30% to each other, can be further grouped into three sub-groups based on i) sequence conservation and, ii) endopeptidase substrate specificity (Drag et al., 2008b; Mukhopadhyay et al., 2006). In a nutshell, SENP1 and SENP2 are thought of as “general” deSUMOylating enzymes with broad specificity towards the three SUMO paralogs, both as endopeptidase and as isopeptidase substrates. SENP3 and SENP5 are SUMO2 and SUMO3 specific and presumably deconjugate nucleolar proteins. SENP6 and SENP7 display poor catalytic rates both on monoSUMO-conjugated proteins and on SUMO precursors and because they are closest to Ulp2 in homology, are thought of as chain-specific deconjugating enzymes. However, this chain-cleaving ability is not unique to SENP6 and SENP7, as it will be revealed in Chapter III of my dissertation.

SENPs as endopeptidases. The first role of SENPs, in the sequence of events during the SUMOylation / deSUMOylation cycle, is the endoproteolytic processing of SUMO precursors to their mature form capable of conjugation. The second role is the isopeptidase activity, and although the biological functions are distinct, the mechanisms are similar. The significant difference between a SUMO precursor and a SUMO-conjugated protein, with respect to the protease, is on the primed side of the scissile bond (Figure 1.3A) When discussing protease substrates one refers to the amino acids found N-terminal to the cleavage site as the unprimed side (P1-Pn), while those C-terminal are referred to as the primed side amino acids (P1’-Pn’). Additionally, the catalytic
domains of SENPs are viewed as having two binding sites for substrates: i) the size-dominant exosite that binds the bulk of the SUMO domain, and ii) the active site groove that binds the C-terminal region. The core of the SUMO molecule positions its C-terminal extension into the SENP active site groove, so that cleavage occurs right after a specific glycine. Both interactions are required for efficient catalysis; SENPs do cleave simple peptides spanning the active site groove, but very slowly. The correct sequence must fit the active site groove, or catalysis is also very inefficient, even in the presence of a SUMO domain (Drag et al., 2008b). Together these sites secure the high specificity and catalytic rates of SENPs for SUMO molecules, and exemplify the power of achieving specificity by the principle of two-site interactions (Scott et al., 2005). The size-dominant exosite brings specificity, but also the risk of product inhibition by the mature form of SUMO precursors (Shen et al., 2006a). However, the biological importance of this inhibition remains to be determined.

The mechanism of SENP specificity is similar to that of DUBs for ubiquitinylated proteins. What differentiates SENPs from DUBs with regards to exosite interactions is the absence of large conformational changes that certain DUBs undergo upon ubiquitin binding. While many DUBs are in a catalytically unfavorable conformation without ubiquitin bound to them (Hu et al., 2002; Renatus et al., 2006), the catalytic domains of SENPs do not seem to undergo large conformational changes upon SUMO-binding and are supposed to constantly be in their active conformation as suggested by the available crystal structures of the catalytic domains of SENP1, SENP2 and Ulp1. Nevertheless,
the lack of such information based on crystal structures should be treated with caution since the enzymes in the crystals conform to whatever state is favored in the crystal lattices and do not necessarily reflect the in vivo state of the proteases. Interestingly, one reported structure suggests a substantial conformational change of SENP8 upon Nedd8 binding (Reverter et al., 2005; Shen et al., 2005). Furthermore, full-length SENPs are yet to be crystallized. Therefore, it should be noted that the lack of conformational change mentioned only applies to the catalytic core of SENPs, and it is formally possible that the large, unstructured N-terminal domains could influence catalysis, however they are likely to have minimal effect on active site architecture itself, therefore I foresee a possible effect on $K_M$, but not on $k_{cat}$ regarding cleavage of SUMO molecules. It should be mentioned that these studies are crucial and highly sought after, however are difficult to achieve based on the low yield of full-length enzymes when produced recombinantly.

There is one more phenomenon exhibited by SENPs that is not explained by the available structural data, which is the ability of SUMO added in trans to enhance SENP activity 10-40 fold on synthetic peptides that occupy only the active site groove (Mikolajczyk et al., 2007). The development of small fluorogenic substrates allowed for the decoupling of the contribution of exosite and active site interactions of SUMO molecules toward SENPs. The mechanism of this substrate-enhanced activity is far from clear, but could in principle be due to slight rearrangements of the active site induced by occupation of the size
dominant SUMO exosite, especially given the minimal conformation adjustments seen upon SUMO binding in currently available structures.

Similar to DUBs, SENPs are intolerant of any residue other than glycine in the P1 and P2 positions (Drag et al., 2008a). This specificity is maintained primarily through a conserved aromatic residue (tryptophane or phenylalanine) that forms a cap over the protease active site prohibiting by steric constraints any amino acid with a side chain. Upstream of these two glycine residues, the other sub-site pockets contribute substantially to SENP cleavage site specificity (Drag et al., 2008b). A study by our lab revealed the expected preference for glutamine in P4 and threonine in P3 for SENP1, SENP2, and SENP5, matching the sequence of the natural SUMO substrate. However, an unexpected preference has beenreveled for SENP6 and SENP7, namely leucine in P4 and arginine in P3. These preferences match the sequence of the cleavage site found in ubiquitin and Nedd8. However, SENP6 and SENP7 cleave neither ubiquitin nor Nedd8 (Drag et al., 2008b), presumably because their overall specificity is defined by exosite interactions. Interestingly, both SENP6 and SENP7 are poor endopeptidases (Mikolajczyk et al., 2007), and since they cluster phylogenetically with yeast Ulp2, they have been suggested to act as polySUMO deconjugating enzymes (Mukhopadhyay et al., 2006). Both SENP6 and SENP7 contain inserts within their catalytic domains that other SENPs lack (Figure 1.5). The insert in the case of SENP6 is dispensable for catalytic function and for SUMO chain specificity (see Chapter III). This is likely also the case for the least investigated member of the family, SENP7, which has also recently
been shown to possess weak endoproteolytic activity (Mikolajczyk et al., 2007) and a more substantial SUMO chain deconjugating activity (Lima and Reverter, 2008). No obvious explanation comes to mind as to why the primary substrate specificity recognition elements in the active site grooves of SENP6 and SENP7 should be so different from the other canonical deSUMOylating enzymes. Interestingly, deletion mutants of SENP7, where possible SUMO chain specificity conferring loops have been deleted from the enzyme, do not show significant changes in their capacity and rate of chain processing, rendering the specificity of chain processing somewhat puzzling (Lima and Reverter, 2008). Chapter III of my dissertation deals extensively with the mechanism of chain deconjugation by human SENPs.

The remarkable specificity of SENPs towards SUMO molecules has lead several groups to develop techniques for the production of recombinant protein fusions in E. coli for the purification of untagged, native proteins through SUMO-fusion systems (Malakhov et al., 2004; Weeks et al., 2007). The protein of choice is fused linearly downstream of Smt3 or SUMO, aiding in expression and solubility of the construct (Marblestone et al., 2006), including membrane-bound proteins (Zuo et al., 2005). The purified SUMO-fusion protein from E. coli is then cleaved with a SUMO protease, to remove the SUMO molecule resulting in the production of the untagged protein of choice after further purification.

**SUMO C-terminal extensions.** As pointed out above, the C-terminal extensions of SUMO molecules are very divergent (Figure 1.1B), and indeed,
the extension of SUMO3 is not even well conserved between closely related animals. Swapping C-terminal tails between SUMO molecules and mutational studies of those tails has provided evidence for the prime side specificity of SENPs (Gong and Yeh, 2006; Mikolajczyk et al., 2007; Xu and Au, 2005). Once SUMO has been conjugated to a target, the role of the C-terminal extension is moot, since it has been removed. A comprehensive study by our lab of the catalytic domains of human SENPs shows different endopeptidase specificities towards the three human SUMO paralogs – summarized in Table 1.1. This specificity is somewhat maintained for SUMO deconjugation (isopeptidase activity) of the model isopeptidase substrate RanGAP1 (Mikolajczyk et al., 2007). This isopeptidase specificity – that the field almost takes for granted – should be treated with caution since it only utilizes a single SUMO-conjugated substrate, RanGAP1. Other SUMO-conjugates could very well have different effects on SENP specificity, especially taking into account possible interactions with the N-terminal domains of SENPs that could contribute to substrate specificity. The challenge of defining SENP isopeptidase specificity will be discussed in detail in Chapter II of my dissertation through the introduction of methods to identify deSUMOylation substrates.

**SENPs as isopeptidases.** SENPs not only participate at the beginning of the SUMOylation/deSUMOylation cycle, but at the very end as well, deconjugating SUMO from their target proteins by cleaving the isopeptide bond between the C-terminal glycine of SUMO and ε-amino group of the lysine side-
Table 1.1 – DeSUMOylating enzymes in a nutshell

1Grouping is based on phylogenetic analysis and substrate specificity determined by a positional scanning library (see text for details). 2S. pombe also has one deSUMOylating enzyme, Nep1. 3Only human SENPs are represented for clarity, however it is generally accepted that other mammalian homologues would have similar characteristics. 4Yeast have only one SUMO paralog, Smt3. 5SENP8 is not a deSUMOylating enzyme, but a deNeddylase, but is a member of the SENP family based on its fold.

<table>
<thead>
<tr>
<th></th>
<th>Structure solved</th>
<th>Preferred Endopeptidase Substrate</th>
<th>Preferred Isopeptidase Substrate</th>
<th>Suggested Subcellular Localization</th>
<th>Specificity Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Yeast</strong> (S. cerevisiae)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ulp1</td>
<td>Catalytic domain only</td>
<td>Smt3 4</td>
<td>Smt3</td>
<td>Nuclear pore complexes (NPCs)</td>
<td>I</td>
</tr>
<tr>
<td>Ulp2</td>
<td>-</td>
<td>?</td>
<td>Smt3 (chain-editing only)</td>
<td>Nucleoplasm</td>
<td>III</td>
</tr>
<tr>
<td><strong>Mammalian</strong> (H. sapiens)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SENP1</td>
<td>Catalytic domain only</td>
<td>SUMO1</td>
<td>SUMO2/3</td>
<td>NPCs</td>
<td>I</td>
</tr>
<tr>
<td>SENP2</td>
<td>Catalytic domain only</td>
<td>SUMO2</td>
<td>SUMO2/3</td>
<td>NPCs</td>
<td>I</td>
</tr>
<tr>
<td>SENP3</td>
<td>-</td>
<td>?</td>
<td>SUMO2/3</td>
<td>Nucleolus</td>
<td>II</td>
</tr>
<tr>
<td>SENP5</td>
<td>-</td>
<td>SUMO2</td>
<td>SUMO2/3</td>
<td>Nucleolus</td>
<td>II</td>
</tr>
<tr>
<td>SENP6</td>
<td>-</td>
<td>Very weak</td>
<td>SUMO2/3 (chain-editing)</td>
<td>Nucleoplasm</td>
<td>III</td>
</tr>
<tr>
<td>SENP7</td>
<td>Catalytic domain only</td>
<td>Very weak</td>
<td>SUMO2/3 (chain-editing)</td>
<td>Nucleoplasm</td>
<td>III</td>
</tr>
<tr>
<td>SENP8</td>
<td>Full-length</td>
<td>Nedd8 only</td>
<td>Nedd8 only</td>
<td>Cytosol</td>
<td>IV</td>
</tr>
</tbody>
</table>

1Grouping is based on phylogenetic analysis and substrate specificity determined by a positional scanning library (see text for details). 2S. pombe also has one deSUMOylating enzyme, Nep1. 3Only human SENPs are represented for clarity, however it is generally accepted that other mammalian homologues would have similar characteristics. 4Yeast have only one SUMO paralog, Smt3. 5SENP8 is not a deSUMOylating enzyme, but a deNeddylase, but is a member of the SENP family based on its fold.
chain of the target protein (Figure 1.3B). All known SENPs – human and others – possess isopeptidase activity when measured on model targets such as RanGAP1 or PML. Choosing a model deSUMOylation substrate allows for the unbiased comparison between the deSUMOylation of different SUMO paralogs (Mikolajczyk et al., 2007), however it does not allow for determination of the isopeptidase specificity of SENPs towards the plethora of different target proteins. The crystal structures of SUMOylated RanGAP1 bound to SENP1 (Shen et al., 2006a) and to SENP2 (Reverter and Lima, 2006) are excellent sets of data showing how the deSUMOylating enzymes bind to substrate-conjugated SUMO molecules, but fall short of providing clues to specific recognition of the SUMOylated substrate by the protease.

Based on catalytic efficiency ($k_{cat}/K_m$ values) of the different SENPs for processing vs. deconjugation, it has been proposed that SENP1 and SENP2 are primarily processing enzymes, whereas SENP3, SENP5, SENP6 and SENP7 are primarily deconjugating enzymes (Mikolajczyk et al., 2007; Mukhopadhyay et al., 2006). This seems all too simple, especially since these assumptions were made based on one or two model deconjugation substrates, therefore far-reaching in vivo conclusions should be avoided. It is quite possible that for certain target proteins SENPs have completely different efficacies. The field has yet to define the relevance of binding partners, cofactors, and the role of the N-terminal domains of SENPs in possibly forming tertiary structures, which in theory can all influence in vivo substrate specificity.
The current explanation for in vivo isopeptidase specificity of SENPs is based solely on putative subcellular localization of these proteases, arguing that deSUMOylation is regulated by the spatial access of the protease to its substrate (Mukhopadhyay and Dasso, 2007). This argument however does not explain the plethora of SUMOylated substrates whose reported subcellular localization does not coincide with that of the proteases (Martin et al., 2007b). Without extensive insight into the role of the N-terminal domains of SENPs – that are currently thought to be responsible for subcellular localization of at least SENP1 (Bailey and O'Hare, 2004), SENP2 (Itahana et al., 2006) and Ulp1 (Li and Hochstrasser, 2003; Sydorskyy et al., 2010) – it is hard to draw conclusions regarding the isopeptidase specificity of SENPs. One way to elucidate the specificity, for example, would be the systematic knock-down of each SENP and determining SUMOylation profiles, a much, needed study, but quite an undertaking. The numerous studies identifying SUMOylation substrates (Panse et al., 2004; Rosas-Acosta et al., 2005b; Vertegaal et al., 2006; Zhao et al., 2004) should be complemented by investigations into deSUMOylation substrates, matching SENPs to SUMOylated substrates.

**SEPNs hiding in the nucleus?** Classically, deSUMOylating enzymes are thought to be associated with the nucleus, with SENP1 (Kim et al., 2005) and SENP2 (Itahana et al., 2006) perhaps shuttling between the nucleus and the cytosol. Another study places SENP2 closely associated with nuclear pore complexes, like SENP1 (Palancade and Doye, 2008). However, which side of
the nuclear pore they actually are on is still a matter of debate. SENP3 and SENP5 have recently been placed in the nucleus, specifically in the nucleolus where SENP3 was shown to be involved in ribosome biogenesis through the deSUMOylation of nucleophosmin (Haindl et al., 2008), while SENP5 has also been implicated in cell division (Di Bacco et al., 2006), and, quite differently, in the maintainence of mitochondrial morphology (Zunino et al., 2007). However, the latter study places SENP5 in the cytosol and is largely based on over-expression studies and suggests a role for SENP5 in deconjugating SUMO1 from target proteins which is not in agreement with substantial data showing SENP5 to be SUMO2/3-specific (Gong and Yeh, 2006). The first study looking at SENP6 (SUSP1 in mice) places it in the cytosol (Kim et al., 2000), but also reports it to be SUMO1 specific, contrary to previous results (Mikolajczyk et al., 2007; Mukhopadhyay et al., 2006). As for SENP7, currently there is no data regarding the whereabouts of the least known member of this family.

SENP regulation is currently explained by the proposed subcellular localization of the proteases, following the notion that catalysis is regulated by keeping the SENPs in a certain compartment, preventing unwanted deSUMOylation. These issues are complicated by experiments showing that a Ulp1 truncation mutant lacking its N-terminal domain rescues Ulp2-null yeast cells from cell cycle arrest, while full-length Ulp1 cannot (Li and Hochstrasser, 2003). This suggests that the N-terminal region of Ulp1 inhibits Ulp1 from deSUMOylating certain substrates. The spatial separation of SENPs and their target SUMOylated proteins may explain why preparation of a cell lysate, when
subcellular compartments are disrupted due to detergents, leads to almost complete deSUMOylation of target proteins. Another possibility is that deSUMOylation is prevented not by masking SUMOylated substrates, but by direct inhibition of SENPs by a hitherto undiscovered SENP inhibitors. To date, no endogenous inhibitors of SENPs have been reported even though cellular inhibitors of various proteases are a common mechanism to regulate their activity (Deveraux et al., 1997). It is only very recently that Rfu1, the first endogenous protein inhibitor of a DUB, has been discovered in yeast, which was shown to regulate the levels of free ubiquitin by inhibiting Doa4 (Kimura et al., 2009). SENP inhibitors could be out there, they just need to be found. On the other hand, the viral protein Gam1 inhibits the SUMO-activating enzyme (E1) by direct binding and completely prevents SUMOylation of host proteins (Boggio et al., 2004), however no such viral inhibitor of deSUMOylation has been described so far. It is not a far-fetched idea to have inhibitors of SENPs around since these proteases are seemingly constitutively active, therefore utilizing endogenous inhibitors would be a very efficient way of preventing unwanted and untimely deSUMOylation.

**SENPs in pathology.** Currently, very little is known of the roles of SENPs in pathology, compared to other intracellular protease systems such as caspases (Salvesen and Riedl, 2008) and DUBs (Reyes-Turcu et al., 2009). The involvement of the SUMOylation/deSUMOylation system in diabetes is an interesting area of research with the GLUT4 glucose transporter reportedly being
SUMOylated (Lalioti et al., 2002), as well as with the negative effect of insulin-stimulated SUMOylation of PTP1B, an ER-membrane-bound protein phosphatase involved in insulin receptor signaling (Dadke et al., 2007). These studies implicate negative regulation by SUMOylation, therefore the action of SENPs would reverse these negative effects to activate the receptors. With SUMO-conjugated substrates involved in a variety of diseases such as cancer and neurodegenerative disorders, it is also understood that SENPs must also play a role in these scenarios, regulating deSUMOylation. In fact, one of the most important tumor suppressor genes, p53, has been show to be positively regulated by SUSP4, a SENP2 homologue from mice, by promoting the self-ubiquitinylating of Mdm2 through its deSUMOylation (Lee et al., 2006). Additionally, p53 itself is also subject to SUMOylation, which is associated with its activation (Gostissa et al., 1999; Rodriguez et al., 1999), a non-traditional effect for SUMOylation of transcription factors.

While several excellent reviews have addressed the role of SUMOylation in cancer (Kim and Baek, 2006; Mabb and Miyamoto, 2007; Seeler et al., 2007), only a few have focused on the role of deSUMOylation (Cheng et al., 2006). For example, SENP1 has been shown to be up-regulated by IL-6 or androgen via the androgen receptor, leading to prostatic intraepithelial neoplasia in mice (Bawa-Khalfe et al., 2007). However, the mechanism through which SENP1 promotes tumor development has not been addressed. SENP1 knock-out mice are available, though the homozygotes are embryonic lethal due to down-regulation of erythropoietin, a HIF1-alpha target gene. The authors show that in
the absence of SENP1, HIF1-alpha remains SUMOylated and is degraded by the proteasome (Cheng et al., 2007), quite an out-of-the-box regulatory mechanism for both HIF1-alpha and for SENP1, leading to cardiac defects in embryonic development. Coincidentally, embryonic cardiac defects are also the cause of lethality of SENP2 deletion in mice (Kang et al., 2010), where SENP2 seems to control polycomb group protein mediated silencing of Gata4 and Gata6. In human genetics, there has been one report of a patient with a fusion of the SENP1 gene to MESDC2. The fusion caused aberrant localization of MESDC2, resulting in infantile sacrococcygeal teratoma, with the product of the fusion gene still exhibiting deSUMOylating activity (Veltman et al., 2005). The SUMOylation system is also perturbed and utilized for infection by certain viruses. Such is the case for the human papilloma virus, whose E1 protein needs to be SUMOylated by host SUMO E3 ligases, such as the PIAS family members, for its nuclear translocation (Rosas-Acosta et al., 2005a). SENPs, in this case, are antagonistic for viral replication, which perhaps suggest that certain viruses could also interfere with the activity of SENPs, although no such observation has been reported. A recent, comprehensive review (Boggio and Chiocca, 2006) delves more extensively into the interplay between viruses and the SUMOylation/deSUMOylation system.

Another group of pathologies where SENPs have been implicated in is a plethora of neurodegenerative diseases, which is summarized in an excellent review by Martin et al. (Martin et al., 2007b). In a nutshell, a rising number of proteins involved in Huntington’s, Parkinson’s and Alzheimer’s disease have
been shown to be SUMOylated, suggesting an unequivocal importance of SENPs in regulating neuronal function. These SUMOylated substrates were shown to be localized not only in the nucleus where they are recruited to neuronal intranuclear inclusions (Takahashi-Fujigasaki et al., 2006), or in the cytosol, such as the SUMOylation of Parkin (Um and Chung, 2006), but even at the plasma membrane, exemplified by the SUMO-modification of the kainaite-receptor (Martin et al., 2007a). The proteases being responsible for their deSUMOylation will surely be determined soon as well, adding a new layer to the complexity of the substrate repertoire and the regulation of SENPs in neurons.

More recently SENPs have been implicated in pathways linked to cancer in general, such as proliferation through the maintenance of proper chromosome alignment through the catalytic activity of SENP6 regulating the SUMOylation of CENP complexes (Mukhopadhyay et al., 2010) or the regulation of DNA repair by SENP6 in controlling the levels of RPA70 SUMOylation (Dou et al., 2010). Furthermore, SENP1 has been implicated over and over again in prostate intraepithelial neoplasia (Bawa-Khalfe et al., 2010). However, in this case it is not clear whether the deSUMOylating activity of the protease is necessary.

**Open questions in SENP biology.** It is quite clear that without the action of SENPs the SUMOylation cycle would come to a halt and the function of transiently SUMOylated proteins would be dramatically affected. A handful of studies have implicated the importance of individual SENPs in deSUMOylating a
handful of substrates both *in vitro* and *in vivo*. However, there is a substantial lack of knowledge on the scope of substrates that are deSUMOylated by SENPs *in vivo* and whether they are conjugated by a single SUMO or by SUMO chains. The recent discovery of SUMO-dependent ubiquitin ligases has further expanded the field and has also underlined the importance of regulated deSUMOylation as carried out by SUMO-specific proteases. Whereas several studies have dealt with the multitude of substrates being SUMOylated, to date there is no study that aims to identify the substrate repertoire of human SENPs on a global scale and determine the degree of overlap between their deconjugating specificity. Therefore, the question remains as to what extent SENPs control the delicate balance between SUMOylation/deSUMOylation, including the deconjugation of substrates modified by SUMO chains.

During my graduate studies I focused on filling in some of the gaps of deSUMOylation, which efforts are summarized in my dissertation. Chapter II uses SENPs as tools in introducing alternative proteomics techniques through proof-of-principle studies that aid in the identification of SUMOylated and polySUMOylated proteins. The techniques are demonstrated through the SUMO system, but are generally applicable to other ubiquitin-like modifiers and their respective proteases. Chapter III, the bulkier part of my dissertation, focuses on mechanistic studies of deSUMOylating enzymes and uncovers the dynamics of deSUMOylation *in vivo*, revealing the abundance of SUMO chains and the way SENPs deconjugate them.
ACKNOWLEDGMENTS

I would like to thank Jowita Mikolajczyk and Marcin Drag for their useful discussions leading up this review. This chapter of my dissertation was reproduced, extended and modified with permission from co-authors from the publication: Békés, M., Salvesen, G.S. (2010). Mechanisms and substrate specificity of deSUMOylating enzymes. In ‘The Ubiquitin Proteasome System in the Central Nervous System: From Physiology to Pathology’ - 2008 update, M. Di Napoli, Wojcik, C., ed. (Nova Science Publishers), pp. 5-32.
REFERENCES


function protease capable of processing the C terminus of Nedd8 and deconjugating hyper-neddylated CUL1. J Biol Chem 278, 28882-28891.


Chapter II

Identifying (de)SUMOylation Substrates by Proteomics
ABSTRACT

The advent of proteomics has allowed the large-scale investigation of components of nearly every biological pathway, and the SUMO system is no exception. With the sophistication of mass spectrometry technologies, the identification of SUMOylated proteins has recently seen a sharp rise, with as many as ~700 SUMOylation substrates identified in humans thus far. However, the fact that a certain protein has been identified as being SUMOylated is only a snapshot of the whole picture, often disregarding the context of deSUMOylation and rarely implies the dynamics of SUMOylation by multiple SUMO molecules. In this chapter of my dissertation, I introduce two proof-of-principle proteomics techniques that define SUMOylation substrates from the protease point of view. One method I developed, dubbed deSUMOMics, is a quantitative proteomics strategy that offers additional information about the substrate regarding the lysine of (de)SUMOylation, while the other, SUMO-protoMAP, is a gel-based proteomics technique that has the potential to identify substrates modified by multiple SUMO entities.
INTRODUCTION

Mass-spectrometry based proteomics (purification/enrichment of the desired proteins, digestion by a protease of choice, LC-MS/MS and database search – in a nutshell) is a powerful tool in identifying specific proteins from complex mixtures (Mann et al., 2001) and has also been widely utilized to determine the post-translational modification of substrates by ubiquitin-like molecules (Mann and Jensen, 2003). For example, the ubiquitinylation and neddylation of a certain protein can be determined by looking for the Gly-Gly stub on lysines in peptides following trypsin digestion (Denis et al., 2007). However, the two modifications are indistinguishable, based on just the stub, therefore more careful data interpretation is advised (Denison et al., 2005). On purified poly-ubiquitinylated proteins, the linkage of ubiquitin chains can also be determined by looking for the Gly-Gly stub on internal ubiquitin-peptides (Kirkpatrick et al., 2005). Other post-translational modifications that have been successfully monitored on a global scale include phosphorylation, although special circumstances during sample preparation are necessary to preserve this labile modification (Rossignol, 2006), and proteolytic signaling deciphered by identifying cleaved substrates (Dean et al., 2007; Gevaert et al., 2003; Mahrus et al., 2008). Specifically, our lab has developed methods to enrich for N-termini from complex mixtures (Timmer et al., 2007), allowing for the identification of the cleaved substrate as well as the cleavage site, which can be profiled on a
proteome-wide scale to determine the specificity of a given protease (Timmer et al., 2009) or a proteolytic pathway, such as cell death (Enoksson et al., 2007).

The identification of SUMOylated substrates is not as straightforward because following trypsin digestion, for example, the SUMO-stub is too large for confident identification of the modified peptides because of overlapping MS/MS spectra stemming from the C-terminus of SUMO and from the modified peptide (Kirkpatrick et al., 2005). However, several bioinformatics strategies have managed to overcome this by “subtracting out” the MS/MS signal from SUMO and thus filtering for SUMO-conjugates (Hsiao et al., 2009; Jeram et al., 2010). So far, these techniques have only been demonstrated on recombinant SUMOylated material and have yet to prove their efficacy when using complex mixtures, such as cell lysates or tissue homogenates. Classically, the identification of SUMOlyated proteins from complex mixtures has been done by stringent purification strategies. In these methods, a tagged SUMO with an exposed C-terminal glycine ready for conjugation is expressed in cells and SUMO-conjugates are isolated by affinity purification, often through multiple steps (Golebiowski et al., 2010), leading to high-confidence hits. However, even when genetically encoded quantification (e.g. SILAC) is employed (Golebiowski et al., 2009), there is a still a significant chance that a certain “SUMOylated” protein is merely a very tight SUMO-interacting protein that has been co-purified with other SUMO-conjugates. Nevertheless, using these techniques many labs have successfully identified proteins modified in yeast by the yeast SUMO, Smt3 (Panse et al., 2004), proteins that are modified by SUMO1 versus SUMO2 in
human cells (Vertegaal et al., 2006), proteins modified by SUMO2 after heatshock treatment (Golebiowski et al., 2009) and have even identified SUMO chains by identifying the internal SUMOylation sites on SUMO2 and SUMO3 (Matic et al., 2008)

Altogether, the current number of “SUMOylated” proteins is upwards of 700 (Golebiowski et al., 2009) and growing. However, *in vivo* validation of most of these hits is still lagging behind, which is understandably a more rigorous undertaking. Whereas, most of the studies focused on the substrates being modified by SUMO, I got interested in utilizing the proteases that deconjugate them to define deSUMOylation substrates. To this end, I developed two proteomics methods to look at deSUMOylation substrates that I will discuss in this chapter of my dissertation.

The first method I introduce, dubbed deSUMOomics, is used to identify deSUMOylation sites by relative quantitation of peptides from a given SUMOylation substrate, while the second one, dubbed SUMO-protoMAP, is a comparative gel-based approach used to identify deSUMOylation substrates based on down-shifts in a gel due to cleavage by SENPs, and could be used to identify substrates modified by multiple SUMO moieties.
RESULTS & DISCUSSION

DeSUMOmics – identifying lysines of deSUMOylation. DeSUMOmics relies on the availability of lysine residues for labeling in untreated and SENP-treated samples. The principle is that the protease will “free” an extra lysine in the cleaved sample, making it available for specific labeling by an amine-reactive reagent. Identifying that lysine by virtue of a specific label will yield the lysine of deSUMOylation, hence SUMOylation. The label of choice in this case is iTRAQ, an isobaric, amine-specific reagent that allows for multiplexing in a single experiment (Zieske, 2006), thereby reducing the number of mass spectrometry runs. The iTRAQ reagents have been designed such that they have identical masses (114-117 Da for 4-plex iTRAQ) but various isotope compositions, so that when iTRAQ-labeled peptides are analyzed in MS1 mode they show up as a single peak because they all have identical mass. However, when the peptide is fragmented in MS2 mode into ‘b’ and ‘y’ ions – to obtain peptide IDs by database searching – the iTRAQ label fragments as well. The reporter peaks in MS2 are unique to a specific label, revealing the extent of modification by a certain iTRAQ reagent given by the intensity of the reporter ion peak. Comparing the intensity of the iTRAQ reporter peaks in MS2 gives an idea of how much a given peptide had been modified by one or more reagents, therefore allowing for relative quantification (Ong et al., 2003).

A general experimental work-flow, as depicted in Figure 2.1, is as follows: 1) prepare SUMOylated substrates, 2) split SUMOylated sample into equal
portions, 3) leave one untreated (control) and cleave the others with SENPs (experimental), 4) denature, reduce and alkylate samples, 5) label differentially with iTRAQ (one iTRAQ reagent per condition), 6) pool samples together in a 1:1 ratio (total protein amount), 7) trypsinize together as a single sample and 8) analyze peptides by LC-MS/MS and database search. I have tested trypsin, chymotrypsin, GluC and AspN for the protease digestion step. Data analysis following LC-MS/MS and database search revealed that AspN digestion resulted in the lowest number of iTRAQ-labeled peptides identified, closely followed by GluC (data not shown). This is probably a combination of the relatively low activity of these proteases and possibly the poor “flyability” of the peptides they generate. Chymotrypsin and trypsin generated roughly the same amount of iTRAQ-labeled peptides, both significantly more than AspN or GluC. Chymotrypsin, however, generated peptides with significantly more multiple internal lysine residues, which are problematic when correctly assigning which lysine bears the iTRAQ label. Trypsin is therefore the ideal protease of choice for iTRAQ-labeled proteins, although it will not cut after iTRAQ-labeled lysine residues and thus the digest could result in slightly longer peptides than it is ideal for proteomics (~25 amino acids). On average, the iTRAQ-labeling efficiency was ~80-90% percent as indicated by the amount of labeled versus unlabeled lysines identified by database search.

In this way, when proteins, not peptides, are labeled by iTRAQ, the native N-termini and lysines (not SUMOylated) will be labeled identically with the different iTRAQ labels, thus the reporter ions coming from the iTRAQ reagents
Figure 2.1 – DeSUMOomics workflow
Schematic representation of the deSUMOomics workflow used to identify deSUMOylation sites. Briefly, SUMOylated proteins are untreated/SENP-treated, then reduced, denatured, alkylated and labeled by iTRAQ and then quenched separately, then pooled and trypsinized together and analyzed by LC-MS/MS, using both CID and HCD fragmentation strategies.
will have equal intensity and will show up with identical peaks in the MS2 spectrum of a given peptide. On the other hand, a lysine that has been “blocked” by SUMO in the control sample, but cleaved off by the SENP in the experimental one, will only show a single peak in MS2 for the iTRAQ signal used for the experimental sample. Therefore, surveying peptides that contain only a single iTRAQ label will identify putative deSUMOylated lysines. Next, database searching of the MS2 spectra with a single iTRAQ will identify the peptide that bears the modification, thus the deSUMOylation site and the protein it belongs to. Classically, iTRAQ analyses have been performed by Q-TOF mass spectrometers. However, with the development of newer generation LTQ machines, such as the OrbiTrap, now they are also capable of analyzing iTRAQ (Kocher et al., 2009). It has to be mentioned that generally LTQ mass-spectrometers utilize CID (collision induced dissociation) to fragment parent ions into ‘b’ and ‘y’ ions for database searching. However, CID is not powerful enough to fragment the iTRAQ label into its reporter fragments, therefore an additional fragmentation step is required, called HCD (high energy collision induced dissociation). Adversely, HCD is too powerful for peptide identification because it blasts the peptide into pieces. Therefore, a conventional way of performing iTRAQ analysis on OrbiTrap instruments is to use both fragmentation energies on a single parent ion, where CID will give the peptide ID and HCD the iTRAQ quantification (Bantscheff et al., 2008). For post-processing after database searching to match CID to HCD data I have used Proteome Discoverer 1.0 from Thermo Scientific for my data analysis.
**DeSUMOomics of recombinant SUMOylated substrates.** In order to demonstrate the basic concept of the method, I prepared SUMOylated substrates in *E. coli* by co-expression of a tri-cistronic plasmid, which encodes the SUMO E1 (Aos1-Uba2), the E2 (Ubc9) and SUMO2 (Uchimura et al., 2004), together with either ∆N-RanGAP1, a canonical SUMOylation substrate (Mahajan et al., 1997), or SUMO2 itself. Due to absence of SUMO isopeptidases in *E. coli*, the co-expression results in high yields of SUMOylated substrates. Besides monoSUMO modification of ∆N-RanGAP1, SUMO2 chains can also modify it, as determined by MALDI analysis. Individual bands were excised from an SDS-PAGE gel of polySUMO2-∆N-RanGAP1 (Figure 2.2A), in-gel trypsinized and then subjected to MALDI-TOF analysis, which showed the presence of SUMO2-SUMO2 isopeptide-linkage specific peptides only in bands modified by two or more SUMO2 entities (Figure 2.2B). When SUMO2 is co-expressed with the SUMOylation machinery, free-floating SUMO2 chains form (Figure 2.3A) that can be purified by classical His-tag purification and can be further separated by anion-exchange chromatography based on the length of the SUMO2 chain (Figure 2.3B). Subjection of the total *E. coli* lysate expressing SUMO2 chains to LC-MS/MS followed by database search using the SUMmOn software, with the help of my collaborators to look for SUMO-conjugates (Jeram et al., 2010), identified the canonical internal SUMOylation site of SUMO2, lysine 11. However, a less abundant linkage at lysine 42 (Figure 2.4) was also identified,
confirming that the SUMOylation system is quite permissive in vitro (Pedrioli et al., 2006).

In a deSUMOomics experiment using lysates from E. coli expressing SUMO2 chains, calculating the ratio of iTRAQ-116 (SENP-cleaved) over iTRAQ-114 (untreated) one can determine whether a given lysine has been deSUMOylated as indicated by an iTRAQ 116/114 ratio that is significantly higher than 1.0, or the mean of the data, if equalization of control vs experimental sample is not perfectly 1:1. This imperfection can be empirically determined and corrected for by plotting the raw intensities of the iTRAQ reporter peaks of a given spectra against each other (Figure 2.5A). After correction by the skew in equalization, plotting the iTRAQ ratios that labeled the free amine of the native N-termini of E.coli proteins in the lysate reveals a Gaussian distribution centered around 1 (Figure 2.5B), suggesting that the labeling efficiency by iTRAQ is comparable in the two samples. On the other hand, the distribution of all iTRAQ ratios (including labeled N-termini and lysines as well) trails to the right (Figure 2.5B), indicating putative deSUMOylation lysines.

However, identifying deSUMOylation lysines of SUMO2 this way is not trivial. In a deSUMOomics experiment using SUMO2 chain expressing E.coli lysates, even if a lysine corresponding to a SUMOylation site of SUMO2 has one of the highest SENP-cleaved over control ratios among the identified SUMO2 peptides, SUMO2 will always have a certain percentage of unoccupied lysines (the internal SUMOylation site on the “top” SUMO2 of a chain), which will skew the ratio, making unequivocal conclusions difficult.
Figure 2.2 – polySUMO2-RanGAP1 is modified by SUMO2 chains
A) 20µg Ni²⁺-purified polySUMO2-RanGAP1-C2 was separated by SDS-PAGE and stained with Coomassie. The asterisk (*) indicates co-purified Ubc9. B) The indicated bands from (A) were excised, in-gel trypsin digested and analyzed by MALDI-TOF mass-spectrometry. The m/z range from 4000-6800 is shown. Arrows indicate the peaks corresponding to the mass of a RanGAP1-C2 peptide modified by the SUMO2-stub (with 0 and 1 missed cleavage in the SUMO C-terminus), black triangles indicate peaks that best match up to the mass of the Lys-11 SUMO2-SUMO2 internal linkage isopeptide (with 0 and 1 missed cleavage in the SUMO2 C-terminus) only found in the bands corresponding to RanGAP1-C2 modified by multiple SUMO2 moieties. Experimental masses were matched to theoretical ones by the PAWS software.
Figure 2.3 – Producing unanchored SUMO2 chains in *E. coli*

A) Lysates producing SUMO2 chains were prepared from *E. coli* (see Experimental Procedures) and were treated with 100nM △N-SENP1 for 1 hour at 37°C to show complete cleavage of SUMO2 chains. Coomassie-stained total lysates are shown on the left, and the corresponding Western blot for SUMO2/3 on the right. The asterisks indicate presumed non-specific bands.

B) Anion-exchange purification of SUMO2 chains on a monoQ column, top panel shows the elution UV trace, bottom panel shows the fractions separated by SDS-PAGE, stained with Coomassie.
Figure 2.4 – Isopeptide linkages of SUMO2 chains
A) The total *E. coli* lysate expressing SUMO2 chains was subjected to trypsin digest and LC-MS/MS and searched using Sequest and SUMmOn algorithms. The top panel shows the total ion current, in the middle the distribution of the parent ion shown for the identified K11-linked SUMO-SUMO adduct and the bottom panel shows the distribution of the parent ion for the K42-linked SUMO-SUMO adduct. B) Left, MS/MS spectra searched by SUMmOn showing SUMO2 modification of SUMO2 at K11, right, MS/MS spectra searched by SUMmOn showing SUMO2 modification of SUMO2 at K42.
Figure 2.5 – iTRAQ-ratios of a deSUMOmics experiment
A) *E.coli* lysate expressing SUMO2 chains was SENP-cleaved and labeled with iTRAQ-116 or left untreated and labeled with iTRAQ-114, analyzed by LC-MS/MS and quantified for iTRAQ reporter peaks. iTRAQ reporter peak intensities for the 116 or the 114 iTRAQ label are plotted per individual spectra. B) Frequency distribution of iTRAQ 116/114 ratios for only labeled N-termini (left) or all quantified iTRAQ ratios, including labeled lysines (right) for the same experiment as above.
Identifying deSUMOylation substrates from cultured cells. To demonstrate the usefulness of deSUMOics in pinpointing deSUMOylation sites from complex mixtures, I established stable cell lines of HEK293A cells expressing mature SUMO1 and SUMO2 (C-terminal Gly) with N-terminal FLAG-tags to aid in purification of SUMOylated proteins. SUMOylated proteins can readily be captured by anti-FLAG immuno-precipitation (data not shown) and can be cleaved with various deSUMOylating enzymes (Figure 2.6A) to remove SUMO-conjugates. Due to the inherent catalytic capacity of the catalytic domains of SENPs (Mikolajczyk et al., 2007), the kinetics of deconjugation on multiple substrates are also different for each SENP.

As a start, I performed deSUMOics on SUMO2-conjugates using ΔN-SENP1 and identified ~200 SUMOylated proteins. Furthermore, 5% of those proteins, contained peptides with lysines that had significantly higher 116/114 ratios, suggesting that they could in fact be deSUMOylation sites (Figure 2.6B). Some examples include the non-POU-domain containing octamer binding protein, elongation factor 1 gamma, spectrin beta and the U5 small nuclear ribonucleoprotein, all of which have been previously shown to be SUMOylated substrates by various proteomics studies (Golebiowski et al., 2009; Vertegaal et al., 2006). However, identifying deSUMOylation sites by differential iTRAQ labeling of lysines relies on there being only a single lysine within a tryptic peptide to which the reporter peak intensities of the iTRAQ labels can be confidently assigned. This somewhat reduces the chances for identifying vast
amounts of deSUMOylation lysines and this is the reason that although during my mass-spec experiments I have identified known SUMOylation substrates, such as RanGAP1 and PML, I was not able to observe the peptides containing their lysines of (de)SUMOylation. Such was the case for TIF-1β and PARP-1 as well, which I have subsequently validated to be true polySUMO2-conjugates in vivo (see below).

As an alternative strategy to profile the deconjugation specificity of SENPs, I have treated SUMO2-conjugates directly on the beads with ∆N-SENP1 and ∆N-SENP6. This way, it is possible to separate SENP-cleaved conjugates (supernatant) from uncleaved material (beads) that remains bound (Figure 2.7A). Analyzing the cleaved material by standard mass spectrometry to identify the deSUMOylated proteins reveals a substantial overlap between ∆N-SENP1 and ∆N-SENP6 (Figure 2.7B) with some putative interesting proteins cleaved by ∆N-SENP6, but not by ∆N-SENP1. This indicates that the catalytic domain of SENPs offers only limited specificity. This is somewhat expected, since SENPs mainly recognize the core SUMO moiety and the catalytic domains presumably make very little contacts with isopeptide-conjugated substrates (Reverter and Lima, 2006; Shen et al., 2006). More interestingly, specificity could be dictated by the catalytic efficacy of the protease, which could also be the point of regulation in vivo, when determining which particular SENP deconjugates a given substrate. Full-length SENPs, in principle, could also dictate alternate isopeptidase specificities, however, such a project would be a whole new study on its own.
Figure 2.6 – deSUMOmics of purified SUMO2-conjugates
A) FLAG-purified SUMO2-conjugates were cleaved with a 2-fold serial dilution of 100nM ∆N-SENP for 30 minutes at 37°C, separated by SDS-PAGE and visualized by anti-FLAG western blotting. B) Frequency distribution of iTRAQ 116/114 ratios of FLAG-SUMO2-conjugates cleaved with ∆N-SENP1.
Figure 2.7 – Isopeptidase specificity of ΔN-SENP1 versus ΔN-SENP6

A) FLAG-SUMO2-conjugates were untreated or cleaved with ΔN-SENP1 or ΔN-SENP6 on the beads, then spun and the supernatant was separated (cleaved substrates) from the still bound uncleaved substrates. Beads and supernatant were separated by SDS-PAGE and the proteins were visualized by SYPRO stain. B) Overlap of deSUMOylated substrates by ΔN-SENP1 and ΔN-SENP6. The supernatants were digested with trypsin and analyzed by LC-MS/MS and database search (Sequest) to identify the cleaved proteins. Non-specific proteins and contaminants were removed from the ΔN-SENP1 and ΔN-SENP6 datasets, peptide probability was set at 0.9, protein probability set at 0.8 with minimum 2 peptides per protein.
Overall, the studies I have conducted on a medium scale identified ~200 SUMOylation substrates, both for the iTRAQ-labeled and unlabeled experiments, that other labs have independently identified as well (Golebiowski et al., 2009). Furthermore, I was able to suggest lysines of deSUMOylation, albeit not as robustly as envisioned, and not necessarily conforming to the SUMOylation consensus. However, a very recent study identified around a hundred SUMOylation sites on a global scale and has uncovered many non-consensus sites (Matic et al., 2010), supporting my findings.

**SUMO-protoMAP identifies polySUMOylated substrates.** Towards the end of my graduate career I became interested again in using mass spectrometry to identify SUMOylated substrates that are modified by multiple SUMO moieties and/or SUMO chains. To this end, I developed a gel-based proteomics strategy that identifies polySUMOylated substrates and differentiates SUMO-interactors from true (de)SUMOylated proteins. The technique, that I termed SUMO-protoMAP, builds largely on the utility of a conjugatable, but not deconjugatable SUMO mutant that I introduce in great detail in Chapter III of my dissertation.

Recently, Dix et al. have developed a proteomics platform called PROTOMAP, which is a technique that visualizes the magnitude of proteolytic events by combined SDS-PAGE/MS/MS (Dix et al., 2008). This method relies on the one-dimensional separation of control protease-treated samples side-by-
side in SDS-PAGE, the systematic division of each lane into multiple gel pieces and the identification of proteins within each band by in-gel trypsin digestion and LC-MS/MS. I adapted this technique to identify polySUMOylated substrates using the aforementioned SUMO2 mutant and by virtue of cleavage by a deSUMOylating protease, which results in the downshift of a SUMOylated protein in SDS-PAGE in the SENP-treated sample. The high specificity of SENPs for SUMOylated adducts allows the specific detection of (de)SUMOylated proteins, and the stringency of SDS-PAGE separation allows the differentiation of true SUMOylated substrates from SUMO-interactors. Ideally, a non-(de)SUMOylated substrate will be identified in gel slices corresponding to its predicted molecular weight and will not shift upon SENP-treatment, therefore it will be in matching gel slices in treated and control samples. On the other hand, a SUMOylated protein will appear in a higher gel slice than where it ought to be in the control sample, but will down-shift to its predicted size upon deSUMOylation in the SENP-treated sample.

As a proof of principle of the method (see Figure 2.8 for the schematic diagram of SUMO-protoMAP), I prepared a cell lysate of the SUMO mutant expressing cells and separated it under native conditions by size-exclusion chromatography. Next, I pooled fractions that contained the HMW-SUMO-conjugates (to enrich for polySUMOylated substrates), split them into two equal parts and treated one with sufficient ΔN-SENP1 to completely remove SUMO-conjugates. After TCA precipitation, I separated the two samples side-by-side by
Figure 2.8 – SUMO-protoMAP workflow
Schematic representation of the SUMO-protoMAP method to identify polySUMOylated substrates. Briefly, HMW-SUMO-conjugates are prepared by size-exclusion chromatography of SUMO2-transfected cell lysates, then untreated or SENP-cleaved, and separated by SDS-PAGE. Each lane in the gel is excised into nine slices and individually in-gel trypsin-digested and followed by LC-MS/MS analysis. The asterisk (*) indicates ∆N-SENP1.
SDS-PAGE, cut each lane into nine 1cm² gel pieces, in-gel trypsin digested all 18 pieces separately, and identified the proteins in each slice by LC-MS/MS and quantified them in a semi-quantitative way by spectral counts (the full list of ID’d proteins per gel slice is available upon request in a sizable excel sheet).

I identified, based on spectral counts, four putative (poly)SUMOylated substrates (including SUMO2 itself) that were present in gel slices higher than their expected size and downshifted upon cleavage by SENP1. SUMO2 (Figure 2.9A) was the only smaller molecular weight protein that was identified to be modified by polySUMOylation. The majority of the proteins identified, such as PKC delta for example (Figure 2.9B), did not down-shift upon SENP treatment, and are therefore unlikely to be modified by SUMO or SUMO chains. However, PARP1 and TIF-1β were identified to be potential (de)polySUMOylation substrates based on their downshift by SENP-treatment (Figure 2.10). They had also been previously reported as SUMOylation substrates by deSUMOomics as well (see above) and by other investigators using mass-spectrometry (Blomster et al., 2010; Golebiowski et al., 2009; Zhao et al., 2004). I confirmed the polySUMOylation of TIF-1β and PARP-1 under endogenous conditions using the SUMO2 mutant (Figure 3.3), demonstrating the power of the SUMO-protoMAP method in identifying endogenous substrates modified by multiple SUMO moieties. Interestingly, PARP1 has recently been shown to be polySUMOylated in Xenopus egg extracts (Ryu et al., 2010) and upon heat-shock in HeLa cells (Martin et al., 2009). Furthermore, oligo-mono-SUMOylation of TIF-1β has been shown to regulate its transcriptional activity (Mascle et al., 2007).
Figure 2.9 – SUMOgraphs of proteins identified by SUMO-protoMAP - I

A) SUMOgraph showing the distribution of SUMO2 – a polySUMO-conjugated protein – in gel slices of untreated (green bars) and SENP-cleaved (red bars) HMW-conjugates. B) SUMOgraph showing the distribution of PKC-delta – a protein unaffected by deSUMOylation – in gel slices of untreated and SENP-treated HMW-conjugates. Corresponding Western blots of the gel filtration fractions used for SUMO-protoMAP are shown for SUMO2 and PKC-delta next to the SUMOgraphs. The x-axis indicates the MW ranges (kDa) of the gel slices, and the y-axis represents combined spectra from duplicate runs on an OrbiTrap-XL mass-spectrometer. The asterisk (*) indicates ΔN-SENP1 cross-reactivity of the antisera.
Figure 2.10 – SUMOgraphs of proteins identified by SUMO-protoMAP - II
Example SUMOgraphs for proteins putatively modified by polySUMOylation, PARP1 and TIF-1β, as determined by spectral counts, showing the down-shift of proteins upon SENP-treatement. The x-axis indicates the MW ranges (kDa) of the gel slices, and the y-axis represents combined spectra from duplicate runs on an OrbiTrap-XL mass-spectrometer. (Figure 3.3 shows the polySUMOylation of PARP-1 and TIFβ in vivo).
Nevertheless, a clear drawback of SUMO/protoMAP – or any other current mass-spectrometry technique looking for SUMO substrates, for that matter – is that it cannot distinguish multiple monoSUMOylation of a substrate from modification by true SUMO chains. Unfortunately, this distinction has to be done on an individual basis by cumbersome mutation of putative lysines to determine where and how many SUMO molecules modify a given substrate. The use of SUMO-SUMO isopeptide-linkage specific antibodies, such as those available for specific ubiquitin-ubiquitin linkages (Matsumoto et al., 2010; Newton et al., 2008), would greatly help in determining the polySUMOylation status of candidate substrates. However, they have yet to be developed. Another way of identifying polySUMOylated substrates is through purifying them by utilizing multiple SIM (SUMO-interacting motif) domains in tandem to enrich for polySUMO-adducts over monoSUMO, a strategy that is quite promising, although it has yet to be demonstrated on a proteome-wide scale.

Out of curiosity, I compared my dataset with that of Golebiowski et al. who identified 766 putative SUMO2 substrates in their heat-shock study (Golebiowski et al., 2009), and found 152 overlapping hits. However, the majority of those proteins did not shift down in my assay upon SENP-treatment. It is possible that these proteins are SUMO2 substrates, but are modified by a single SUMO2 molecule, resulting in a ~10kDa shift in mobility in SDS-PAGE which would be missed because of the resolution of the SUMO/protoMAP technique. On the other hand, SUMO/protoMAP affords a much more stringent
analysis of SUMOylation because it distinguishes SUMO conjugates from SUMO-interactors, which could have skewed the heat-shock study.

Furthermore, SUMO-protoMAP not only identified polySUMOylated proteins, but it also identified K48-linkage ubiquitin peptides only in the higher gel slices, and not in the lower ones (Figure 2.11A), indicating that the HMW fractions used for the assay could contain polyubiquitinylated proteins as well. Indeed, a small number of proteins I identified were in gel slices higher then their expected size, but did not down-shift upon SENP-treatment, suggesting that those could be modified by ubiquitin- rather than SUMO2 chains. Due to the nature of the SUMO2 mutant, that can be conjugated but not removed (Q90P), it is quite possible that besides polySUMO-modified substrates, free-floating SUMO chains were also present in the HMW fractions. Interestingly, performing a transfection experiment with a ubiquitin construct that carries the analogous mutation in its cleavage site (L73P), results in a beautiful laddering of various-length ubiquitin chains, clearly demonstrating the formation of free-floating ubiquitin chains (Figure 2.11B) in vivo. This, in principle, supports the idea that free-floating SUMO chains should be able to form in a similar fashion in vivo, however they have yet to be formally established.

Therefore, the versatility of SUMO-protoMAP is quite apparent, as it could easily be adjusted to investigate the polymeric nature of other ubiquitin-like modifiers and in combination with isopeptide-linkage specific antibodies could in the future be the ultimate method in deciphering signaling by chains of ubiquitin-like modifiers.
Figure 2.11 – The utility of SUMO-protoMAP in visualizing ubiquitin chains
A) Total ion chromatogram (top, black) and extracted ion chromatograms for an internal ubiquitin peptide (green) or the K48 ubiquitin-ubiquitin isopeptide (burgundy) shown for four gel slices. B) Wild-type or the L73P ubiquitin mutant with an N-terminal HA-tag was transfected into HEK293A cells, lysed under native conditions in the presence or absence of NEM, separated by SDS-PAGE and visualized by the indicated antibodies.
MATERIALS & METHODS

Preparation of SUMO-conjugated recombinant substrates. E. coli-generated polySUMOylated proteins were prepared by co-transformation of BL-21 cells with the desired substrate and with the pT-E1E2S2 plasmids (a gift of Prof. Hisatoh Saitoh) and expressions were induced with 0.2mM IPTG, followed by incubation for 6 hours at 30°C. E. coli cells were sonicated in 50mM Tris, pH 8.0 and 100mM NaCl and cleared by centrifugation and passing through a 0.45µM filter. 250µg of the total E. coli lysate containing polySUMO2-chains was cleaved with 100nM ΔN-SENP1 to show collapse of SUMO2 chains. All His-tagged proteins were purified by Ni²⁺-NTA purification and step-eluted with 20-200mM imidazole in 50mM Tris, pH 8.0, 150mM NaCl. To further purify polySUMO2, purified chains were run on a MonoQ (5/5) anion exchange column in bis-Tris, pH=6.0 and eluted by 5-300mM NaCl using an Ekta FPLC system.

MALDI analysis of polySUMO2-conjugated RanGAP-1. 20µg of purified, dialyzed and concentrated polySUMO2-conjugated RanGAP-1-C2 (from X. leavis) was run on an 8-18% ammediol gel, stained with Coomassie-stain and the indicated individual bands were excised and subjected to in-gel tryptic digestion as described previously (Dix et al., 2008). Extracted peptides were resuspended in 50% aceto-nitrile, 0.1% TFA and mixed 1:1 either with saturated sinapinninc acid (SA) or CHCA solution in 50% aceto-nitrile, 0.1% TFA and
analyzed by a MALDI-TOF instrument (Bruker Daltonics). Experimental masses were matched up with theoretical ones using the PAWS software (http://bioinformatics.genomicsolutions.com/Paws.html).

**DeSUMOmics of recombinant SUMOylated substrates.** For iTRAQ-labeling experiments, 4mg of total *E. coli* lysate expressing SUMO2 chains or 400µg of purified SUMO2 chains or 400µg of polySUMO2-RanGAP were either left untreated or cleaved with 100nM of ∆N-SENP1 for 1 hour at 37°C in 50mM HEPES, pH=7.4, 150mM NaCl in the presence of 10mM DTT and 2% was run on SDS-PAGE and stained by Commassie to confirm cleavage by the protease. The two samples (uncleaved/cleaved) were separately denatured by 6M guanidine by incubation at 60°C for 1 hour, then treated with 30mM iodoacetamide (IAA) for 45 minutes at room temperature in the dark, then buffer exchanged using a PD-10 column into freshly prepared, deionized 8M urea, 50mM HEPES, pH=7.4 and precipitated overnight by TCA/acetone precipitation. Samples were then resuspended in iTRAQ Dissolution Buffer and 100µg material was labeled by iTRAQ-114 (uncleaved) and by iTRAQ-116 (SENP-cleaved) rocking at room temperature for 1 hour, excess label was quenched by 50mM glycine, pH=7.4 for 30 minutes at room temperature. Next, samples were mixed 1:1 and cleaved by trypsin (or chymotrypsin, AspN or GluC) together overnight at 37°C, dried down, C18-ZipTip-cleaned the next day and resuspended in 2% acetonitrile, 0.1% TFA.
DeSUMOmics of endogenous SUMO substrates. Stable cell lines expressing FLAG-SUMO1 and FLAG-SUMO2 were originally created by transfecting 5µg plasmid into HEK293A cells, then the cells were selected for 3 weeks with 0.6mg/ml G418 and individual clones expressing near-wild-type levels of SUMO were picked and continuously maintained in full DMEM supplemented with 0.2 mg/ml G418 to maintain a minimal selection pressure. Stable cells were grown up in 10x15cm dishes and harvested in 50mM HEPES, pH=7.4, 150mM KCl, 1mM EDTA, 1% NP-40, 0.1% CHAPS, supplemented with protease inhibitors E-64, DCl, leupeptine, MG-132 and in the presence of 30mM N-ethylmaleimide (NEM, Pierce). To purify SUMO-conjugates lysates were incubated with anti-FLAG agarose beads overnight at 4°C, washed twice in lysis buffer and twice in PBS and either kept on beads at 4°C or eluted with 3X FLAG-peptide. SUMO-conjugates were then cleaved immediately with the indicated ∆N-SENP(s) (100nM) and prepared for iTRAQ labeling the same way as described above for recombinant proteins. For all other protein identification of SENP-cleaved substrates, the same protocol was applied, except the iTRAQ labeling step.

LC-MS/MS and database search. Samples were desalted on C_{18}-ZipTip columns, redissolved in 0.1% trifluoroacetic acid and 2% acetonitrile and 1-3µg material was loaded onto an OrbiTrap-X_L mass spectrometer (Thermo) via a C_{18}-column (0.2 x 150mm Michrom Magic, 3µM particles, 200Å pore size) and eluted on a 180min RP gradient (Brill et al., 2009) with the MS/MS method being a top six, data dependent method. Precursor ions were scanned from m/z 300-
2000 in the Orbitrap at a resolution of 30,000, then the three most abundant precursor ions were fragmented in the HCD cell at 50% collision energy, and the resulting fragment ions, with a first m/z ratio of 100, were scanned in the Orbitrap at a resolution of 15,000. A threshold of 1e5 for MS/MS is specified, and dynamic exclusion is enabled for 60 s repeat duration and 60 s exclusion duration with a repeat count of 2. Subsequently, the same three precursor ions were fragmented by CID at the usual 35% energy and scanned in the linear ion trap at a resolution of 60,000. Peptides were identified from MS/MS spectra by the Sequest algorithm using Proteome Discoverer 1.0 (Thermo Scientific, CA) by searches against the appropriate protein database and reverse database, with a precursor mass tolerance of 10.0 ppm, specific for static carbamidomethylation of Cys residues and differential iTRAQ labeling of N-termini and lysine residues, differential oxidation of Met residues, and full-enzyme specificity. Proteome Discoverer automatically assigns peptide ID obtained through CID and iTRAQ quantification by HCD, however some manual filtering of the data is required to sort out duplicate peptide IDs and wrongfully assigned iTRAQ labels. Confidence of hits was set to 5% FDR, singletons were included for the iTRAQ quantification, however a minimum of three spectral count was applied to standard protein ID.

**SUMO-PROTOMAP analysis.** Two 10cm-dishes of HA-SUMO2-Q90P transfected HEK293A cells were lysed, as before, in the presence of 30mM NEM and separated on a Superose-6 size exclusion column. Fractions were
analyzed by SDS-PAGE and Western blotting to identify HMW-SUMO-conjugates, which fractions were pooled and differentially cleaved by 0.5µM ΔN-SENP1 for 6 hours at 37°C. Next, untreated and SENP-treated fractions were TCA precipitated, analyzed by Western blotting to ensure complete cleavage of SUMO-conjugates by SENP1 and separated by SDS-PAGE on an 8-18% ammediol gel. The untreated and SENP-treated lane was cut into nine equal 1cm² gel pieces and all 18 slices were individually in-gel trypsinized and C_{18}-ZipTip cleaned. The 18 gel slices were analyzed in duplicates by LC-MS/MS on an OrbiTrap mass-spectrometer using a 1-hour RP separation followed by fragmentation in CID of the top 4 data-dependent MS ions. MS/MS spectra were searched using Sequest and analyzed and compiled by Proteome Discoverer 1.1. Identified proteins are visualized individually in a “SUMOgraph” by plotting spectral counts (combined from the two technical replicates) against the different gel slices they were found in. All proteins identified (singletons included, minimum 3 spectral counts, with an FDR of less than 5% for peptide ID) are listed in Table 2.1.

**Ubiquitin L73P experiments.** The Leu to Pro mutation in the P4 position of the cleavage site of ubiquitin (–LRGG) was introduced by standard PCR. Wild-type and the L73P mutant was cloned into pcDNA3 with an N-terminal HA-tag using HindIII-XhoI restriction sites. Plasmids were transfected into HEK293A cells using NanoJuice and cells were lysed, 48 hours after transfection, separated by SDS-PAGE and visualized by anti-ubiquitin and anti-HA antibodies.
ACKNOWLEDGEMENTS

I would like to specifically acknowledge the help of Laurence M. Brill in contributing to this chapter of my dissertation by patiently running all my samples on the mass-spec and providing excellent feedback and insight in fine-tuning the OrbiTrap to correctly blast the peptides using HCD so that the iTRAQ signal could be quantified properly. I would also like to thank John Timmer and Mari Enokkson for keeping me on track with doing proteomics, Majid Ghassemian and Khatereh Motamechaboki for assisting with iTRAQ and MALDI experiments, respectively, and Tharan Srikumar and Brian Raught who carried out the SUMmOn search.
REFERENCES


Chapter III

The Mechanism and Dynamics of SUMO Chain Deconjugation by

Human SUMO-specific Proteases
ABSTRACT

SUMOylation of proteins is a cyclic process that requires both conjugation and deconjugation of SUMO moieties. Besides modification by a single SUMO, SUMO chains have also been observed, yet the dynamics of SUMO conjugation/deconjugation remain poorly understood. Using a non-deconjugatable form of SUMO, I establish that SUMO chains are abundant in vivo, I highlight the importance of SUMO deconjugation and demonstrate the highly dynamic nature of the SUMO system. I show that SUMO-specific proteases (SENPs) play a crucial role in the dynamics of SUMO chains in vivo by constant deconjugation. Preventing deSUMOylation in Schizosaccharomyces pombe results in slow growth and a sensitivity to replication stress, highlighting the biological requirement for SUMO dynamics. Furthermore, I present the mechanism of SUMO chain deconjugation by SENPs, which occurs via a stochastic mechanism, resulting in cleavage anywhere within a chain. Our results offer mechanistic insight into the working of deSUMOylating proteases and highlight their importance in the homeostasis of (poly)SUMO-modified substrates.
INTRODUCTION

Reversible post-translational modification of proteins by ubiquitin-like covalent modifiers is a widely utilized mechanism to alter the fate, binding partners, function or localization of a given target protein (Schwartz and Hochstrasser, 2003). A prime example is the covalent attachment of multiple ubiquitin moieties to a lysine side-chain of a target protein (Hershko, 1983), leading to the formation of ubiquitin chains, which have different functions depending on the lysines utilized in the internal ubiquitin linkage. Recently, small ubiquitin-like modifier (SUMO) has also been shown to form chains in vitro (Knipscheer et al., 2007; Tatham et al., 2001) and in vivo (Matic et al., 2008; Schimmel et al., 2008).

The SUMO conjugation pathway consists of SUMO activation, transfer, and a ligation enzyme machinery analogous to the ubiquitin pathway (Johnson et al., 1997; Lois and Lima, 2005; Reverter and Lima, 2005). Somewhere upwards of 500 cellular proteins are SUMOylated in vivo (Golebiowski et al., 2009), yet the extent of observable SUMOylation is only a snapshot due to the presence of SUMO-specific proteases, SENPs. The SUMO cycle begins and ends with specific proteolytic events: the processing of proSUMO and the deconjugation of SUMO from the target protein by SENPs (Drag and Salvesen, 2008). In addition to monoSUMOylation, there is now growing evidence that SUMO, like ubiquitin, forms polymeric chains. Thus, the non-covalent interaction of the SUMO conjugating enzyme, Ubc9, with SUMO has presented a possible
mechanism for SUMO chain formation (Capili and Lima, 2007; Knipscheer et al., 2007) and indeed in vitro, all SUMO molecules (Smt3, the S. cerevisiae SUMO homologue, and human SUMO1, -2 and -3) have been observed to form SUMO chains, primarily via Lys residues located near their N-termini (Geiss-Friedlander and Melchior, 2007). In vitro the SUMOylation system is quite permissive, allowing for multiple SUMO-SUMO linkages (Jeram et al., 2010). The initial indication that SUMO chains exist in vivo came from transfection experiments yielding multimers of SUMO2 conjugated to HDAC4 (Tatham et al., 2001) and from Saccaromyces cerevisiae (Bylebyl et al., 2003). More recently, evidence for in vivo SUMO chains comes from mass-spectrometry experiments using HeLa cells, revealing chain formation via Lys-11 of SUMO2 and SUMO3 (Matic et al., 2008). However, the extent of polySUMO vs. monoSUMO modification of the proteome has yet to be established, although polySUMOylation has been reported to increase during heat-shock treatment (Golebiowski et al., 2009).

Once SUMO chains have been assembled on a substrate they could have two fates. They can either be disassembled by SENPs, presumably recycling the SUMO monomers and the target protein (Drag and Salvesen, 2008; Hay, 2007) or, at least in theory, delivered to the proteasome via SUMO-targeted ubiquitin E3 ligases (STUbls) that cause the ubiquitin-mediated degradation of polySUMO-modified substrates (Perry et al., 2008).

SENPs are cysteine proteases that belong to peptidase clan C48 and share a common fold of their catalytic domain (Li and Hochstrasser, 1999). In S. cerevisiae there are two deSUMOylating enzymes, Ulp1 and Ulp2, while
humans have six (SENP1, -2, -3, -5, -6 and -7). SENPs are dual function proteases that can remove the C-terminal extensions from SUMO precursors and deconjugate SUMO from modified proteins (Békés, 2010). They exhibit substantial substrate specificity (Mikolajczyk et al., 2007; Shen et al., 2006) and a distinction has been established that sets SENP6 and SENP7 apart from other human SENPs, largely because of their poor ability to process SUMO precursors and monoSUMOylated RanGAP1. Rather, SENP6 and SENP7 have been proposed to be chain-specific, partly because SENP6 and SENP7 are most similar to yeast Ulp2, which has been shown to limit the amount Smt3 chains in yeast (Bylebyl et al., 2003).

Several questions remain to be answered - how are SUMO chains recognized and deconjugated by SENPs, how dynamic is the process, and how important are the dynamics? To answer these questions I utilized two human SENPs, SENP1 and SENP6, that diverge substantially in specificity and sequence conservation to determine the mechanism of SUMO chain deconjugation. Furthermore, by using a conjugatable, but not deconjugatable version of SUMO, which targets physiological substrates, I investigated how SENPs participate in and regulate the dynamics of SUMO chains in vivo. We also investigated the influence of irreversible SUMOylation on cellular growth rate and sensitivity to replication stress in S. pombe.
RESULTS

The dynamics of SUMO chain deconjugation by endogenous SENPs.

When cells are lysed under non-denaturing conditions almost no SUMOylated proteins are observable. However, including a cysteine modifier such as N-ethyl maleimide (NEM) or a denaturant, such as SDS, during cell lysis retains a certain level of SUMOylation (Golebiowski et al., 2010), and the general conclusion is that endogenous constitutively active SENPs rapidly deconjugate SUMOylated proteins as a lysis artifact. I hypothesized that SENPs continuously deconjugate SUMO from its targets in vivo, and that a dynamic balance exists between SUMO addition and deconjugation, which is not represented in the instantaneous steady state captured in cell lysates. To address the dynamics of the SUMOylation cycle in vivo, I utilized a mutation within the cleavage site of SUMO that allows for conjugation, but is defective in deconjugation by SENPs (Owerbach et al., 2005).

I created eukaryotic expression constructs of wild-type (wt) mature SUMO2 and the aforementioned Q90P mutant, containing N-terminal epitope tags, and transfected them into HEK293A cells. The conjugating enzymes of the SUMO pathway tolerate the Q90P mutant because both the SUMO E1 and E2 have large active site cavities (Olsen et al., 2010; Reverter and Lima, 2005) and can accommodate a slightly kinked SUMO C-terminus introduced by Pro in P4 (Figure 3.1), whereas the narrow active site cleft of SENPs does not tolerate it (Drag et al., 2008). Upon cell lysis, Q90P SUMO2 remains extensively
Figure 3.1 – Enzymes of the SUMO pathway modeled with a P4 Gln to Pro (Q90P) SUMO2 mutant
The C-terminus of SUMO2 with either Gln (wild-type) or Pro (Q90P) in P4 are visualized sitting in the active sites of the E1 enzyme (pdb: 3KYD), Ubc9 (1Z5S) and ∆N-SENP-1 (2IYD), showing how stearic clashes with the protease only could prevent cleavage of the Q90P SUMO2 mutant by SENPs but could allow conjugation by the E1-E2 cascade. The images were generated in Pymol.
conjugated in even in the absence of NEM or SDS (Figure 3.2), demonstrating that while the SUMO conjugation machinery accepts this mutant, endogenous SENPs are unable to remove it during lysis and presumably in vivo as well. To confirm the validity of the Q90P mutant in "permanently" SUMOylating known endogenous substrates, I looked for SUMO-conjugates whose multi-SUMOylation status can be observed under endogenous conditions using the Q90P mutant. I have confirmed the polySUMO modification of PARP-1 and TIF-1β (Figure 3.3), SUMOylated proteins that I have previously identified by two alternative mass-spectrometry techniques presented in Chapter II.

Further demonstrating the rapid dynamics of the SUMO pathway, in a transfection time-course experiment showing the expression of wild-type and the Q90P mutant SUMO2 (Figure 3.4), both monomeric unconjugated and higher-molecular-weight (HMW) conjugated species of wild-type SUMO2 can be seen as early as six hours post-transfection. Importantly, the Q90P mutant mainly appears as HMW conjugates with very little observable unconjugated monomer. This suggests that wild-type SUMO2 continuously cycles in a very dynamic process – presumably due to the action of endogenous SENPs, which fail to deconjugate the Q90P mutant, leading to an accumulation of stably SUMOylated proteins.

To investigate the nature of the HMW-conjugated SUMO-species, I enriched them by size exclusion chromatography (Figure 3.5) and cleaved them with serial dilutions of the catalytic domain of SENP1 (ΔN-SENP1). SUMO-conjugates from wild-type SUMO2 HMW fractions were cleaved very rapidly
Figure 3.2 – SUMO2 Q90P is not deconjugated by endogenous SENPs

A) HA-tagged wild-type and Q90P SUMO2 were transfected into HEK293A cells and lysed after 24 hours in the presence or absence of 30mM NEM or 2% SDS. Lysates were separated by SDS-PAGE and visualized by Western blotting for the HA-tag, demonstrating cleavage resistance of the Q90P mutant. Two dilutions of the same lysates are shown.
Figure 3.3 – SUMO2 Q90P targets endogenous substrates

HEK293A cells were co-transfected with wild type SUMO2 (HA-tag) and wild-type SUMO3 (myc-tag), or with Q90P SUMO2 (HA-tag) and Q90P SUMO3 (myc-tag), together with simultaneous knock-down of SUMO2 and of SUMO3 by siRNA, treated with a single dose of 50 kJ/m² UV-irradiation, allowed to recover for 5 hours at 37°C, and total cell lysates were analyzed by the indicated antibodies. Dotted lines indicate cropping within the same gel.
Figure 3.4 – SUMO2 Q90P reveals the dynamics of SUMOylation \textit{in vivo}

Lysates (+NEM) from a time-course expression analysis of HEK293A cells transfected with either HA-tagged wild-type or Q90P SUMO2 constructs were separated by SDS-PAGE and visualized by Western blotting for the HA-tag, revealing the differential conjugation extent of the Q90P mutant. Dotted line indicates cropping within the same gel.
Figure 3.5 – Purification of HMW-SUMO2-conjugates
HEK293A cells expressing either HA-SUMO2-WT (left) or HA-SUMO2-Q90P (right) were lysed in the presence of 30mM NEM and the lysates were separated on a Superose-6 size exclusion column. UV-trace of lysates is shown on the top and Western blot for the HA-tag of the fractions indicated is shown on the bottom, displaying a chain-like banding pattern of SUMO-conjugates.
(EC$_{50}$ in the low nM range) and reaction products were converted exclusively to mono-SUMO moieties with no observable intermediates (Figure 3.6A). In contrast, cleavage of HMW SUMO-conjugates from the Q90P SUMO2 mutant shows very different kinetics (Figure 3.6B). Quantifying the disappearance of HMW conjugates demonstrates that the Q90P SUMO2 mutant is eventually deconjugatable with very high levels of \(\Delta N\)-SENP1, but with substantially decreased kinetics – a 1000-fold change in EC$_{50}$ (Figure 3.7A). This suggests that while the deconjugation reaction is possible \emph{in vitro}, it is highly unlikely to happen \emph{in vivo}. More importantly, the reduced rate of cleavage now reveals cleavage intermediates (Figure 3.6B). The intermediates observed are consistent with variable length SUMO chains generated due to cleavage by SENPs. Importantly, a SUMO2 Q90P mutant, wherein the three Lys residues in the N-terminus (K6, K8 and K11) were mutated to Arg to limit chain elongation (Tatham et al., 2001), prevented the appearance of multimeric cleavage intermediates (Figure 3.7B). Thus, by definition and logic, the universe of SUMOylated proteins contains abundant SUMO chains.

In control experiments I demonstrate that the HMW-conjugates are indeed polySUMO adducts, because they are cleaved by SENPs but not by a promiscuous deubiquitinating enzyme (Figure 3.8A). I ruled out the possibility that a putative endogenous \textit{E. coli}-purified factor could be responsible for the deconjugation because a catalytically dead \(\Delta N\)-SENP1 with an Ala replacing the active site Cys does not remove the HMW SUMO-conjugates (Figure 3.8B) and I also show that SUMO chains of variable lengths from numerous gel-filtration
Figure 3.6 – HMW-SUMO2-conjugates in the Q90P SUMO2 fractions consist of polySUMO2 chains - I

HMW fractions from lysates of HEK293A cells expressing wild-type (A) or Q90P mutant (B) SUMO2 were obtained by size-exclusion chromatography and the fractions were treated with 3-fold serial dilutions of 1µM ∆N-SENP1, separated by SDS-PAGE and visualized by Western-blotting for the HA-tag. Multimeric cleavage intermediates are indicated on the right.
Figure 3.7 – HMW-SUMO2-conjugates in the Q90P SUMO2 fractions consist of polySUMO2 chains - II

A) Quantification of cleavage of wild-type and Q90P HMW-conjugates from a duplicate experiment demonstrates that the Q90P SUMO2 is eventually cleaved, but with a 1000-fold increase in EC₅₀. B) A parallel experiment is shown using ΔN-SENP1 to cleave HMW-conjugates from N-all-Arg-SUMO2-Q90P expressing cells, notice the absence of multimeric cleavage intermediates.
Figure 3.8 – Cleavage of HMW-SUMO2-conjugates by isopeptidases
A) FLAG-SUMO2-Q90P HMW-conjugates were prepared as described and were cleaved with 0.5µM of the indicated His-tagged SENPs (ΔN-SENP1, -2, -5, -6, -6-Δloop, -7) or DUBs (3 – UCH-L3, 7 – USP-7CD) for 30 minutes at 37°C and visualized by SDS-PAGE and Western blotting for FLAG (the substrate) and for His (the protease). Dotted lines indicate cropping within the same gel.

B) HA-SUMO2 HMW-conjugates were prepared as described and cleaved with 100nM of wild-type or catalytically dead (active site cysteine mutant) ΔN-SENP1 for 30 minutes at 37°C and visualized by SDS-PAGE and Western blotting for the indicated antibodies. Note, that even the active ΔN-SENP1 does not remove polyubiquitin from the HMW-conjugates.
Figure 3.9 – Cleavage of various length polySUMO2 chains

FLAG-SUMO2-Q90P HMW-conjugates were prepared as described and different fractions from the gel filtration (#21, #27 and #31) containing different length SUMO2 chains were cleaved with 10-fold serial dilutions of 1µM ∆N-SENP1 and visualized by SDS-PAGE and Western blotting for the FLAG-tag.
Figure 3.10 – Full-length SENP1 also cleaves SUMO2 Q90P HMW-conjugates

N-terminally FLAG-tagged full-length and ΔN-SENP1 were purified from HEK293A cells and used to cleave HA-SUMO2-Q90P HMW-conjugates in a time-course assay for the indicated times at 37°C. The cleavage reactions were visualized by SDS-PAGE and Western blotting for the HA-tag, revealing that full-length SENP1 cleaves SUMO chains as well.
fractions are collapsed in a similar fashion (Figure 3.9). Notably, the N-terminal
domain of SENP1 does not substantially influence cleavage of Q90P HMW-
SUMO-conjugates because full length, wild type SENP1 also trims SUMO
chains to SUMO-multimers (Figure 3.10), showing that SUMO chain
decoupling is an intrinsic property of human full-length SENP1.

**Free and substrate-anchored SUMO chains are substrates of SENPs.**

Having demonstrated the abundance of SUMO chains in vivo, I wanted to know
the reason behind their presumed ephemeral existence. An obvious explanation
is that they are such excellent substrates for SENPs that the cleavage
intermediates are hard to demonstrate experimentally, since they are only
observed when using the Q90P mutant. To demonstrate that SENPs cleave
SUMO chains, I purified SUMO chains expressed in *E. coli*, either as free-
floating chains or anchored to the canonical SUMO substrate RanGAP1, by
employing a polycistronic expression vector encoding SUMO E1, E2 and
SUMO2 (Uchimura et al., 2004), as previously (see also Chapter II).

The catalytic domains of both SENP1 and SENP6 (ΔN-SENP1 and ΔN-
SENP6), as examples of the different classes of SENPs (Békés, 2010),
efficiently deconjugate SUMO2 chains produced in *E. coli*, whether or not they
are attached to a heterologous protein acceptor, such as RanGAP1 (Figure
3.11). Interestingly, even though our recombinant enzymes are comparably
reactive towards an activity-based probe (data not shown), ΔN-SENP1 cleaves
SUMO2 chains more efficiently than does ΔN-SENP6. It can therefore be
Figure 3.11 – Cleavage of recombinant SUMO chains by SENPs

Purified unanchored SUMO2 chains (A) or purified polySUMO2-RanGAP1-C2 (B) were cleaved with serial dilutions of ∆N-SENP1 and ∆N-SENP6 for 30 minutes at 37°C then separated by SDS-PAGE and visualized by Western blotting for SUMO2/3 or stained with Coomassie stain, demonstrating the extent of SUMO2 chain deconjugation.
concluded that SUMO chain deconjugation – at least in vitro – is not a unique property of SENP6 and SENP7, but that other SENPs share this activity.

Preliminary studies that I conducted with heterogeneous mixtures of multiple SUMO chains of various lengths did not allow us to determine whether SUMO chains are cleaved i) from the free chain end, one SUMO at a time (processive deconjugation) or ii) randomly anywhere within the chain (stochastic deconjugation) to give rise to the various SUMO multimer intermediates. I therefore developed an in vitro model system to specifically address the mechanism of SUMO chain deconjugation.

**Linear SUMO chains are cleaved in a stochastic manner.** I designed a set of linear tri-SUMO2 constructs mimicking isopeptide-linked chains to generate a homogenous system in which I can address the importance of each cleavage site within a chain individually. SUMO chains are predominantly linked by Lys-11 in vivo (Jeram et al., 2009; Matic et al., 2008). To closely approximate this linkage, I generated linear concatamers lacking the first 11 residues of SUMO2. In these linear constructs, three ΔN11-SUMO2 molecules are fused “head-to-tail” (N- to C-terminal by a peptide bond) to simulate an isopeptide-linked tri-SUMO2 (Figure 3.12A). Each concatamer has an N-terminal FLAG-tag and a C-terminal His-tag that allows discrimination of the fragments after SENP-cleavage based on the size of the cleaved SUMO monomers (Figure 3.12B) and by Western blotting (Figure 3.12C). Next, I used the cleavage-prohibitive mutation Q90P to direct cleavage by SENPs at specific sites within the linear
Figure 3.12 – Linear tri-SUMO2 chains as a mechanistic model
A) Schematic representation of linear tri-SUMO2 constructs. Open green circles (○) represent cleavable (WT) ∆N11-SUMO2 and full red circles (●) represent uncleavable (Q90P) ∆N11-SUMO2. The lighting bolt (♂) indicates the cleavage site. B) Purified tri-SUMO2 proteins were diluted to 1µM and cleaved with 50nM ∆N-SENP1 for 30 minutes at 37°C and visualized by SDS-PAGE and Coomassie-staining. All tri-SUMO2 proteins are only cleaved at the allowed wild-type cleavage site (where P4 is Gln) and not when P4 is mutated to Pro (Q90P). C) Western blot for the FLAG-tag (N-terminus) and for the His-tag (C-terminus) of the tri-SUMO2 proteins, showing the expected cleavage fragments and the correct orientation of the various tri-SUMO2 cleavage mutants.
Figure 3.13 – Cleavage of tri-SUMO2 constructs by ΔN-SENP1

2μM tri-SUMO2 with Pro in all three cleavage sites (A) or tri-SUMO2 with two allowed cleavage sites (B) or a tri-SUMO2 with one cleavage site (C) or proSUMO2 (D) were treated with a 2-fold serial dilution of 1μM ΔN-SENP1, showing that the triple Pro-containing tri-SUMO2 (three full red circles) is not cleaved at all, while the double-cleavable tri-SUMO2 (green-green-red circles) is cleaved simultaneously at both cleavage sites, indicated by the appearance of all three monomeric SUMO moieties, also the cleavage of the tri-SUMO2 is comparable to proSUMO maturation. The cleavage assays were visualized by SDS-PAGE and coomassie staining. The black triangles (▼) indicate approximate EC\textsubscript{50} values.
Figure 3.14 – Cleavage of tri-SUMO2 constructs by SENPs is stochastic
The indicated tri-SUMO2 proteins (2µM) were cleaved with 2-fold serial dilutions of 1µM ΔN-SENP1 (A) or ΔN-SENP6 (B) for 30 minutes at 37°C and visualized by SDS-PAGE and Coomassie-staining. The order of cleavage sites within tri-SUMO2 for each panel is indicated in the bottom right corner by the order of the red and green circles. “E” stands for lane with enzyme alone and “S” stands for lane with substrate only. Panels in B have been cropped for clarity. The black triangles (▼) indicate approximate EC50 values.
Figure 3.15 – The loop within the catalytic domain of SENP6 is dispensable for activity and specificity
A) Wild-type or ∆loop (see Figure 1.5) SENP6 was labeled with the active site probe HA-SUMO2-VME in a time-course, reactions were run on SDS-PAGE and visualized by anti-His western blotting. B) 2µM tri-SUMO2 with cleavage allowed after the first SUMO moiety (left) or the last (right) was treated with a 2-fold serial dilution of 1µM ∆N-SENP6-∆loop, showing that, similar to wt ∆N-SENP6 (Figure 3.14), the loopless mutant also does not discriminate between cleavage sites.
Figure 3.16 – Full-length SENP1 cleaves tri-SUMO2 constructs

2µM tri-SUMO2 with cleavage allowed after the first SUMO bead (top) or after the last (bottom) was treated with a 2-fold serial dilution of purified FLAG-tagged full-length SENP1 (purified from HEK293A cells). The graph on the bottom shows the quantification of the cleavage by quantifying the uncleaved triSUMO2 band in each lane and expressed as percentage of the no enzyme (control) lane. The asterisks (*) indicate full-length SENP1.
chain (Figure 3.12A). If all three available cleavage sites contain the Q90P mutation, the tri-SUMO2 construct is uncleavable (Figure 3.12B and 3.13A) in the context of the linear constructs, where SENPs act as endopeptidases.

I defined EC\(_{50}\) as the amount of protease required to cleave 50% of the substrate in a given time. For those constructs containing only a single cleavage site, there was no significant difference in the EC\(_{50}\) values, irrespective of the position of the cleavage site within the constructs for either ∆N-SENP1 (Figure 3.14A) or for ∆N-SENP6 (Figure 3.14B). Thus, there is no preferential recognition of any SUMO moiety within such a linear chain. This is further supported by allowing cleavage after both the first and second moiety, resulting in the simultaneous processing of the tri-SUMO2 construct from both ends, giving rise to the appearance of all three SUMO monomers (Figure 3.13B). However, I noticed that cleavage after the last SUMO moiety, just before the His-tag, slightly enhanced cleavage (2-fold decrease in EC\(_{50}\)) for all the SENPs tested. This is consistent with their preference for shorter C-terminal fragments when assayed on proSUMOs (Mikolajczyk et al., 2007). Furthermore, the cleavage of a triSUMO2 construct is comparable to proSUMO2 maturation (Figure 3.13C-D), suggesting that other than the small preference of SENPs for short C-terminal fragments, the geometry of the triSUMO2 does not significantly influence cleavage rates. For SENP6 it has been previously suggested that an insertion loop within its catalytic domain could be responsible for its polySUMO-specificity (Lima and Reverter, 2008), yet deletion of the loop had no significant effect on its activity (Figure 3.15A) nor did it influence its preference for either
cleavage site when assayed on triSUMO2 (Figure 3.15B), arguing against the proposal and suggesting that the core catalytic domain of a SENP is sufficient for SUMO chain deconjugation.

The data point to a stochastic mechanism for chain cleavage of the model tri-SUMO2 constructs, and reveal that both SENP1 and SENP6, as examples of different classes of SENPs, share this mechanism. Therefore, specificity is directed by the primary recognition mechanism of a single SUMO moiety, at least at the level of the catalytic domain of SENPs. Full-length human SENP1 also does not discriminate between cleavage sites when assayed on linear triSUMO2 constructs (Figure 3.16), indicating that mechanistically it is no different from its catalytic domain. Taken into account that full-length SENP1 also trims HMW-conjugates of SUMO Q90P (Figure 3.10) it is safe to say that whatever the function of the N-terminal domain of SENP1 might be, it has no significant restrictive effects on the processing of SUMO chains.

**Cellular consequences of the lack of SUMOylation/deSUMOylation.**

In the fission yeast *S. pombe* deletion of Pmt3 (the sole *S. pombe* SUMO homologue) displays a phenotype with an increased sensitivity to DNA damage (Tanaka et al., 1999). To address the biological role of deSUMOylation my collaborators generated Pmt3-Q108P to replicate the Q90P mutation of human SUMO2. GFP-tagged wild-type Pmt3 and the Q108P mutant were ectopically expressed in the pmt3Δ strain, and as expected, the Q108P mutant enhanced stable Pmt3-conjugates (Figure 3.17A). Additionally, when plated on
hydroxyurea, △Pmt3 cells expressing the Q108P mutant grow just as slowly as the deletion strain (Figure 3.17B). This demonstrates that not only SUMOylation, but also deSUMOylation is required for an efficient response to replicative stress. These findings are in line with the phenotype of S. cerevisiae lacking Ulp2, where this deSUMOylating protease is required for cell division following genotoxic stress (Schwartz et al., 2007).

To perform similar tests in human cells, I sought to deplete HEK293A cells of both SUMO2 and SUMO3 simultaneously and reconstitute them with various SUMO mutants. The depletion of both SUMO2 and SUMO3 by siRNA resulted in an aberrant cellular morphology leading to large, swollen, conjoined cells (Figure 3.18). Reconstitution of double knock-down HEK293A cells with wild-type HA-tagged SUMO2 and myc-tagged SUMO3 simultaneously was able to partially rescue the aberrant morphology, while double transfection using the Q90P mutants failed to do so (Figure 3.19). Double knock-down HEK293A cells reconstituted with the Q90P SUMO mutants exhibited slower growth, indicating that in the absence of endogenous SUMO-2/3 the non-deconjugatable mutant is disadvantageous for cellular growth (Figure 3.20). Whereas it has previously been shown that SENPs are important in cell function, these results place the need for deSUMOylation and – by the definition and example through the Q90P mutant – for SUMO chain deconjugation in an evolutionary context, highlighting the extensive dynamics of (de)SUMOylation across lower and higher eukaryotes.
Figure 3.17 – *In vivo* effects of the lack of deSUMOylation in *S. pombe*

(A) Western blot analysis of the indicated *S. pombe* strains was performed using denaturing lysis conditions to analyze total levels of SUMOylated proteins *in vivo*. Cells were grown in minimal media, that either permitted (-B1: Left panel), or repressed gene expression (+B1: Right panel). The asterisk indicates a non-specific band. (B) Five-fold dilutions of the indicated strains were grown in minimal media, which either repressed (+B1), or de-repressed expression (-B1) of the indicated genes, in either the presence, or absence of the replication inhibitor hydroxyurea.
Figure 3.18 – Simultaneous knock-down of SUMO2 & SUMO3 in HEK293A cells
Phase-contrast images showing HEK293A cells left untreated, treated with non-specific siRNA, or a pool of siRNA against SUMO2 and SUMO3. Notice the aberrant morphology of some of the SUMO-2/3 silenced cells (orange triangles). Scale bar is 50µm. The bottom right panel shows Western blots against SUMO-2/3 and Hsp-90 (as loading control) of lysates prepared from the cells shown in the images.
Figure 3.19 – Simultaneous knock-down of SUMO2 & SUMO3 and concomitant reconstitution by wild-type and Q90P rescue plasmids

A) Phase-contrast images of HEK293A cells 48 hours after transfection with non-specific or SUMO2 and SUMO3 siRNAs and concomitant reconstitution with wild-type or Q90P SUMO2 and SUMO3. Scale bar is 100µm. B) Corresponding Western blots showing knock-down efficiency. Notice how reconstitution with the Q90P plasmids does not rescue the phenotype of SUMO-2/3 knock-down.
Figure 3.20 – In the absence of endogenous SUMOs, SUMO2 Q90P results in a slow-growth phenotype.

Growth curve (top) of double knocked-down and reconstituted HEK293A cells. Cells were grown in 24-well plates and counted every day for 3 days post-transfection by Trypan Blue exclusion (the error bars represent SEM of the experiment in duplicate). Western blot analysis (bottom) of the cells shown in panel A. “ns” and “S” stand for non-specific and SUMO-2+3 siRNA, respectively. Asterisks (*) indicate non-specific bands.
DISCUSSION

The mechanism of SUMO chain deconjugation by SENPs. I demonstrate that SUMO chain cleavage by the catalytic domains of human SENP1 and SENP6 is a very rapid and stochastic event. Clearly I cannot rule out that SENPs not investigated in my dissertation may have a processive ability. However, I observed no mechanistic difference between the catalytic domain of SENP1 and SENP6 (the most divergent human SENPs), and so I propose that stochastic SUMO chain cleavage is a universal mechanism across the SENP family because of the common fold of these proteases. This should also apply to non-human deSUMOylating enzymes such as S. cerevisiae Ulp1, which is structurally similar to human SENPs (Mossessova and Lima, 2000), and has also been shown to cleave in vitro generated SUMO1 chains (Pedrioli et al., 2006). It has been proposed that insertion loops in the catalytic domains of SENP6 and SENP7 could enhance their polySUMO specificity, however the deletion of the SENP7 loop did not impair its ability to cleave SUMO chains (Lima and Reverter, 2008). Furthermore, the insertion loop of SENP6 is also dispensable for its activity, specificity and its chain cleaving mechanism. This is in line with the results presented here that the core catalytic domain of SENPs alone, such as ΔN-SENP1, can catalyze SUMO chain deconjugation. Therefore, the difference between SENPs is not in terms of the mechanism of how they recognize polySUMO, but in terms of rates of cleavage of SUMO-conjugates. It
has also been recently suggested that the N-terminal domains of SENPs, which could contain putative SUMO-interacting motifs (SIMs), may confer a processive chain-cleaving specificity (Kolli et al., 2010). However, I observed the same stochastic mechanism of chain cleavage using both full-length and catalytic domain only forms of SENP1 – arguing against this proposal, at least for SENP1.

My studies conducted on linear substrates as a model of endogenous chains point to a quite flexible arrangements of the SUMO moieties within tri-SUMO2, which is reflected in modeling isopeptide-linked SUMO2-SUMO2 linkages (Figure 3.21), suggesting that a ∆11-SUMO2 trimer could be a faithful mimic of isopeptide-linked SUMO chains. Furthermore supporting a stochastic chain-cleaving mechanism, of isopeptide-linked conjugates as well, is the capture of multimeric cleavage intermediates when cleaving SUMO Q90P HMW-conjugates. Thus, the synthetic linear chains are valuable surrogates for natural SUMO polymers. As emerging crystal structures of various di-ubiquitin linkages have revealed different conformations of ubiquitin chains (Bremm et al., 2010; Virdee et al., 2010), I eagerly await atomic level structural studies on di-SUMO molecules that would reveal the extent of flexibility of endogenous SUMO-chains.

A unified stochastic mechanism of SUMO chain deconjugation for SENPs is in contrast to ubiquitin chain deconjugation by several DUBs, which have evolved different substrate recognition mechanisms and ancillary ubiquitin binding domains, giving rise to DUBs that prefer different ubiquitin linkages or
Figure 3.21 - Modeling isopeptide-linked SUMO2 chains
Modeling a 5-mer SUMO2 chain based on RanGAP-SUMO linkage from the crystal structure of RanGAP1-SUMO1:SENP2 (pdb: 2IO2) shows how the protease (in ribbon structure with transparent surface in pink and light brown) has sufficient binding capability at any given SUMO-SUMO isopeptide linkage (surface representation of linked SUMO2 molecules in shades of blue). The image was generated using Pymol.
various length isopeptide-linked ubiquitin (Komander et al., 2009) and others that cleave only mono-ubiquitin (Reyes-Turcu et al., 2009). Importantly, many DUBs, undergo significant conformational changes upon substrate binding (Hu et al., 2002; Reyes-Turcu et al., 2006; Reyes-Turcu et al., 2009), but SENPs show no such rearrangements within their catalytic domains (Reverter and Lima, 2004; Shen et al., 2006) and are presumed to be constitutively active.

SENPs display a great degree of specificity for the three human SUMO paralogs in terms of SUMO-precursor activation and in terms of the deconjugation of monoSUMOylated RanGAP1, the canonical model substrate that has been used to assay the isopeptidase activity of SENPs (Kolli et al., 2010; Mikolajczyk et al., 2007; Shen et al., 2006). This specificity is largely based on contacts at the SUMO-SENP interface, with minimal conserved contacts elsewhere. However, there is a kinetic difference in the cleavage of peptide-linked versus isopeptide-linked SUMO moieties, such that SENP1 is a better isopeptidase than an endopeptidase for SUMO2 (Mikolajczyk et al., 2007). The isopeptide bond between SUMO and its target is better positioned for cleavage than the equivalent peptide bond in proSUMO, as revealed in the crystal structures of the RanGAP1-SUMO2:SENP2 complex compared with proSUMO2:SENP2 (Reverter and Lima, 2006). This supports my findings that the Q90P SUMO2 mutant in a linear construct (peptide-linked) is less sensitive to cleavage than in natural conjugates (isopeptide-linked). Thus, although I see absolutely no cleavage of the peptide-linked Q90P SUMO2, I observe low, but
measurable, activity of SENP1 on isopeptide-linked Q90P SUMO2 HMW-conjugates.

**DeSUMOylation in vivo.** The number of biological pathways that SUMOylation has been shown to be involved in is expanding (Makhnevych et al., 2009), but it is not clear how important deSUMOylation is in these events. Our results, in which the phenotype of the \( \Delta \text{Pmt3} \) *S. pombe* sensitivity to replication stress is maintained upon reconstitution with the Gln to Pro mutant, demonstrate the importance of deSUMOylation *in vivo*. Along this line and by definition, the deSUMOylation of SUMO chains is just as important. Even though the Q90P mutant strongly suggests unbalanced SUMO chain formation, a biological function for SUMO chain formation is equivocal. An Smt3 mutant unable to form SUMO chains rescues some of the phenotypes caused by the deletion of the deSUMOylating enzyme Ulp2, with the conclusion that SUMO chain formation is not required for essential functions in *S. cerevisiae* (Bylebyl et al., 2003), although SUMO chains clearly form *in vivo*.

Silencing of SUMO2 and SUMO3 in U2OS cells leads to a modest sensitivity to heat-shock treatment (Golebiowski et al., 2009), therefore it was surprising to uncover the phenotype of silencing both SUMO2 and SUMO3 in HEK293A cells, where a subset of cells displayed aberrant morphologies. This phenotype could not be rescued by reconstitution with uncleavable (Q90P) SUMO2 and SUMO3 mutants, and is similar to the phenotype observed upon SENP6 depletion in HeLa cells that leads to aberrant kinetochores.
(Mukhopadhyay et al., 2010). Interestingly, another human deSUMOylating enzyme, SENP5, has also been shown to be required for cell division (Di Bacco et al., 2006), further arguing for the importance of deSUMOylation, which, as I have shown, includes dynamic regulation of SUMO chains.

Genetic knock-outs or cellular knock-downs of deSUMOylating enzymes can highlight the requirement for the protein itself, but unless the knock-down is accompanied by reconstitution with a catalytic mutant (Kang et al., 2010; Li and Hochstrasser, 1999; Mukhopadhyay et al., 2010) one cannot determine whether the catalytic activity or just the presence of the protein is required. The application of cleavage-resistant mutants takes an orthogonal approach in defining the importance of substrate cleavage on a more global scale, and has been successfully deployed to define the role of specific proteolysis of proteins such as coagulation factor V in regulating coagulation (Dahlback, 1995), and retinoblastoma protein, in uncoupling its tumor promoting function from its tumor suppressor one (Borges et al., 2005). SUMO2 Q90P fulfills this role because it demonstrates the importance of deSUMOylation on a global scale and allowed me to define the mechanism of SUMO chain deconjugation. The inability of SENPs to remove SUMO Q90P leads to a massive shift in SUMOylation dynamics, and uncovers the prevalence of SUMO chains in vivo. Thus, my studies offer an important additional insight into the biology of human SENPs by highlighting their role as active negative regulators of SUMO chains, a feat that would have been difficult to achieve globally using genetic manipulation techniques.
The dynamics of SUMO chains. The intrinsic and rapid chain deconjugation activity of human SENPs renders SUMO chains a challenging phenomenon to observe and decipher. It is generally assumed that tight spatial and temporal regulation of encounter between SENPs and their substrates is required in vivo. For example, Ulp1 in S. cerevisiae is sequestered in the nucleolus following stress by low-levels of alcohol, allowing transient SUMO conjugate build-up (Sydorskyy et al., 2010). However, my results imply that a subset of SENPs constantly cleaves SUMOylated proteins and SUMO chains, essentially contributing an editing function that balances the ligation of SUMO onto its targets. This hypothesis is supported in principle by the finding that RanGAP1 is SUMOylated equally well by both SUMO1 and SUMO2, but in vivo it is conjugated to SUMO1, because the SUMO2-conjugated RanGAP1 is constantly deSUMOylated by SENPs (Zhu et al., 2009).

Importantly, the multimeric cleavage intermediates only observed with the Q90P SUMO2 mutant reveal that a portion of HMW-conjugated species are modified by SUMO chains in vivo, and suggest that SUMO chains could be such excellent substrates for SENPs that under normal experimental conditions they are difficult to detect. It follows that the majority of SUMO-conjugated proteins identified (Golebiowski et al., 2009; Hsiao et al., 2009; Vertegaal et al., 2006) may even represent the worst possible SENP substrates. These are the proteins that “escape” deSUMOylation and are detectable by various experimental
conditions that can only prevent deSUMOylation upon cell lysis and not during the life of the cell.

The SUMO Q90P mutant identified physiologically relevant multi-SUMOylated substrates, and I suggest that it could also lead to the formation of free, unanchored SUMO chains, which could be a heterogeneous mixture of endogenous wild-type and transfected Q90P SUMO moieties. The idea of free SUMO chains having a signaling capacity has recently surfaced (Ulrich, 2008), however it has yet to be experimentally demonstrated, although the results of SUMO-protoMAP discussed in Chapter II point in this direction. On the other hand, unanchored polyubiquitin was recently shown to have a role in activating the kinase TAK1 (Xia et al., 2009). I suggest that an important function of SENPs is to monitor SUMO chain formation, whether conjugated or free-chains, to ensure that polySUMO-interactors function appropriately (Lallemand-Breitenbach et al., 2008; Tatham et al., 2008). An emerging concept in cell signaling proposes a fundamental role for tight regulation of ubiquitin chain editing, driven by the deconjugating proteases, in pathways such as NFκB signaling (Sun, 2010) and death receptors signaling (Wertz and Dixit, 2010). Based on my observations, I propose a parallel function for SUMO chain dynamic regulation through SENPs in cell signaling. I also envision that the utility of conjugatable, but not deconjugatable ubiquitin-like modifiers will aid in tackling the physiological and pathological consequences of regulation by this highly versatile form of post-translational modification.
MATERIALS & METHODS

Size exclusion chromatography of cell lysates and Western blotting. SUMO-transfected HEK293A lysates were sedimented twice at 13,200rpm in a 4°C table-top centrifuge to clear insoluble material. The lysates were separated by size-exclusion chromatography on a 24mL Superose-6 gel-filtration column (GE Healthcare) in 20mM Tris, pH 7.6, 150mM NaCl and 1mM EDTA using an Akta FPLC system. Collected fractions (0.5mL) were stored at -20°C overnight or cleaved immediately with a serial dilution of the indicated proteases in the presence of 5mM DTT for 30 minutes at 37°C and analyzed by SDS-PAGE and Western blotting. Primary antibodies were detected by classical chemiluminescence using HRP-conjugated secondary antibodies (GE Healthcare) or by fluorophore-coupled secondary antibodies (LiCOR) and quantified using the Odyssey system. Quantification was based on the uncleaved material, expressed in percentages compared to the no-enzyme control lane.

In vitro deSUMOylation assays. DeSUMOylation reactions of E. coli-purified SUMO-conjugates were carried out in 50mM Tris, pH 8.0, 20mM NaCl and 10mM DTT. 4 µg of SUMO2-chain-conjugates were cleaved with 2-fold serial dilutions of 100nM SENPs as indicated for 30 minutes at 37°C, the reactions were terminated by boiling in SDS-loading buffer. Results were visualized by Coomassie-staining or Western blotting with anti-SUMO-2/3 antibodies.
Design and cleavage of tri-SUMO2 constructs. Three repeats of ∆N11-SUMO2 were linked in a linear fashion “head-to-tail” one linkage at a time by ligation following individual PCR reactions and finally cloned into pET28 with an N-terminal FLAG-tag and a C-terminal His-tag. The SENP-cleavage site at the P4 position (Schechter and Berger, 1967) was alternated between Gln (wild-type) and Pro (Q90P mutant). The three beads were amplified independently by PCR with the following restriction sites: Ncol-AgeI (bead #1), AgeI-Tsp45I (bead #2) and Tsp45I-Xhol (bead #3). The full list of primers for this intricate cloning is available upon request. For cleavage assays, tri-SUMO2 constructs were diluted to 2µM final concentration and cleaved with a 2-fold serial dilution of 1µM SENPs for 30 minutes at 37°C, then visualized by SDS-PAGE and Coomassie-staining. All constructs were verified to have the correct orientation of cleavable (wild-type) or non-cleavable (Q90P) tri-SUMO2 beads by DNA sequencing as well as Western blotting for FLAG and for His following +/- SENP-treatment of the purified proteins.

Cell culture, transfection, lysate preparation and purification. HEK293A and HeLa cells were readily maintained in Dulbecco’s modified Eagle’s medium (DMEM, Cellgro) supplemented with 10% heat-inactivated fetal bovine serum, 100units/mL penicillin, 100µg/mL streptomycin and 2mM L-glutamine. Full-length and catalytic-domain SENP constructs and all the SUMO-variants were transfected into cells using NanoJuice (Novagen) according to the
manufacturer's protocol. 6-48 hours post-transfection cells were washed twice with cold PBS and lysed on ice in 50mM HEPES, pH=7.4, 150mM KCl, 1mM EDTA, 1% NP-40, 0.1% CHAPS, supplemented with protease inhibitors E-64, DCl, leupeptine and MG-132 in the presence or absence of 30mM N-ethylmaleimide (NEM, Pierce), alternatively 2% SDS was included in the same lysis buffer where indicated and after two quick PBS washes at room temperature, the cells were boiled directly. To purify FLAG-tagged SENPs from HEK293A cells, lysates were incubated overnight at 4ºC with anti-FLAG agarose-beads (M2, Sigma), then washed three times in lysis buffer, and twice in PBS and eluted with 3X-FLAG-peptide (Sigma) and labeled with HA-SUMO2-VME for 30min at 37ºC to ensure their activity. Lysates were used immediately for subsequent experiments or kept at -20ºC.

**Knockdown of SUMO2 and SUMO3 by siRNA.** A pool of five individual siRNA oligos (Qiagen) that target the 3'-UTR region of the SUMO2 and the SUMO3 mRNA were used in the experiments at 3nM final concentration per individual siRNA. HEK293A cells were reverse-transfected with the siRNA alone or together with rescue plasmids according to the manufacturer’s protocol and the cells were harvested 48 hours post-transfections and the lysates were probed with the indicated antibodies. The phase-contrast images of cells were taken with a camera attached to an Olympus IX71 microscope (Olympus) at 100X and 200X magnification. For growth curve experiments HEK293A cells were reverse-transfected at low confluence in 6-well plates and immediately transferred to 24-
well plates, split six-ways and maintained in full DMEM for the indicated times, then trypsinized and counted manually using a hemocytometer by Trypan Blue (Sigma) exclusion.

**S. pombe experiments.** pREP41-eGFP-pmt3-Q108P was cloned via standard PCR protocols using pmt3 cDNA and primers described in Supplemental Methods. The resulting product was then Ndel-BamHI digested and cloned into pREP41. Western blot analysis of total levels of SUMOylation has been described previously (Prudden et al., 2007). The *S. pombe* strains used for this study are listed in Table 3.1.

**Plasmid constructs.** The cloning of SUMO proteins and SENP catalytic domains has been described previously (Mikolajczyk et al., 2007). The constructs for the catalytic domain of *Xenopus laevis* RanGAP1 (RanGAP1-C2) and the pT-E1E2S2 plasmid were generous gifts of Prof. H. Saitoh. Full length SENPs were cloned from a keratinocyte library and cloned into pcDNA3.1 vectors with an N-terminal FLAG-tag (J. Mikolajczyk), using EcoRV-XhoI restriction sites. For mammalian expression, SUMO2 proteins (UniProtKB Swiss-Prot, P61956) were subcloned from pET28 vectors into pcDNA3.1 with N-terminal FLAG- or HA-tags and with a mature C-terminus ending in Gly, followed by a stop codon, using EcoRV-XhoI (FLAG-tag) or HindIII-XhoI (HA-tag) restriction sites. SUMO3 (UniProtKB/Swiss-Prot P55854) was subcloned from pET28 into pcDNA3.1 with an N-terminal myc-tag using restriction sites HindIII-
XhoI. The C-terminal glutamine to proline mutation in P4 was introduced by PCR amplification using a primer containing the mutation. The loop-3 deletion mutant of ∆N-SENP6 (UniProtKB/Swiss-Prot Q9GZR1) was created by deleting amino acids 773-925 by over-lapping PCR and cloned into pET29 with a C-terminal His-tag by Ndel-Xhol restriction sites. All constructs were verified by sequencing and all the primers used are available upon request.

**Protein expression and purification using E. coli.** All recombinant proteins were produced in *E. coli* BL-21 cells (Novagen), unless stated otherwise. The expression and purification of SENP catalytic domains has been described previously (Mikolajczyk et al., 2007). The expression of SUMO-constructs, including tri-SUMO2 mutants, pro-SUMOs, and RanGAP1-C2 were induced with 0.2mM IPTG (isopropyl-β-D-galactopyranoside) and expressed for 3 hours at 37°C. *E. coli*-generated polySUMOylated proteins were prepared by co-transformation of BL-21 cells with the desired substrate and with the pT-E1E2S2 plasmids and expressions were induced with 0.2mM IPTG, followed by incubation for 6 hours at 30°C. *E. coli* cells were sonicated in 50mM Tris, pH 8.0 and 100mM NaCl and cleared by centrifugation and passing through a 0.45μM filter. 250μg of the total *E. coli* lysate containing polySUMO2-chains was cleaved with 100nM ∆N-SENP1 to show collapse of SUMO2 chains. All His-tagged proteins were purified by Ni²⁺-NTA purification and step-eluted with 20-200mM imidazole in 50mM Tris, pH 8.0, 150mM NaCl. The concentrations of purified proteins were determined by absorbance at 280nm based on molar absorption
coefficients and calculated by the Edelhoch relationship (Edelhoch, 1967). Protein purities were determined by Coomassie-staining. All samples were used immediately for further experiments or were stored at -80ºC. USP-7_{CD} was purchased from BostonBiochem.

**Antibodies used.** Mouse anti-His (34660, Qiagen), rabbit anti-SUMO2/3 (polyclonal, in-house raised against purified SUMO2 as the immunogen, referred to as anti-SUMO-2/3 because it recognizes both SUMO2 and SUMO3), mouse anti-FLAG (F-3165, Sigma), mouse anti-HA (16B12, Covance), mouse anti-myc (sc-40, Santa Cruz Biotechnology), mouse anti-PARP (556362, BD Pharmingen), mouse anti-CAND-1 (sc-137124, Santa Cruz Biotechnology), rabbit anti-ubiquitin (polyclonal, in-house produced), rabbit anti-cyclin A1 (sc-751, Santa Cruz Biotechnology), mouse anti-Hsp90 (610418, BD Transduction Laboratories), mouse anti-Hsp60 (SPA-806, Stressgen), rabbit anti-cleaved-PARP (9541S, Cell Signaling), rabbit anti-TIF-1β (4123S, Cell Signaling), anti-pmt3 (polyclonal, in-house produced), rabbit anti-protein kinase C delta (sc-937, Santa Cruz Biotechnology).

**Sequences of siRNA against SUMO2 and SUMO3.** SUMO2 target sequences: ‘5-TAGCTGTACATGTAGGGCAA-3’ & ‘5-CTGTCTTTAAGTAGGGATAAA-3’ & ‘5- GTCAAGCGTCTTGTTTAA-3’ and SUMO3 target sequences: ‘5- AGGGATGAATCTGTAACTTAA-3’ & ‘5- TTGGGATCTATACCAATTAAA-3’.
**Fission yeast strains used in this study.** All strains are *ura4-D18, leu1-32*. NBY2761, *pmt3::ura4*⁴ pREP41-EGFP-*pmt3-Q108P:LEU2*⁺; NBY2773, *pmt3::ura4*⁴ pREP41-*blank:LEU2*⁺; NBY2774, *pmt3::ura4*⁴ pREP41-EGFP-*pmt3:LEU2*⁺.

**Primers used to clone *S. pombe* Pmt3-Q108P (Ndel-BamHI).**

(cgcgcccatatgctgaatcaccatcagcaaac),

(ggggcgccgatcttaaccactaaagttctaatagcc).
ACKNOWLEDGEMENTS

I would like to thank D. Fasci who has done an excellent job in preparing the HA-SUMO2-VME activity-based probes and M. Drag, Z. Ronai, S. Riedl, T. Hunter, S. Reed, R. Hampton, M. Petroski and B. J. van Raam for stimulating discussion, M. Enoksson, L. M. Brill and K. Motamedchaboki for help with LC-MS/MS, and C. Carell and P. Mace for SUMO modeling. M.B. designed, performed and interpreted the experiments with the following exceptions, Tharan Srikumar carried out the SUMmOn search in Brian Raught.’s laboratory, and John Prudden and Michael N. Boddy performed the S. pombe studies. M.B. and G.S.S. conceived the project and wrote the manuscript, which all authors have read and agreed to. The authors declare no conflicts of interest. This chapter of my dissertation reproduced and modified with permission from co-authors from the manuscript: Békés, M., Prudden, J., Srikumar, T., Raught, B., Boddy, M.N., and Salvesen, G.S., The mechanism and dynamics of SUMO chain deconjugation by human SUMO-specific proteases, submitted for publication in 2010.
REFERENCES


high accuracy mass spectrometry and an in vitro to in vivo strategy. Mol Cell Proteomics 7, 132-144.


deconjugation regulates paralog-selective sumoylation of RanGAP1. Mol Cell 33, 570-580.
Chapter IV

Conclusions & Perspectives
OUTLOOK ON THE WORLD OF DESUMOYLATION

In this concluding chapter of my dissertation I would like to put my results of SENP isopeptidase specificity (Chapter II) and deSUMOylation dynamics (Chapter III) into the broader context of SENP biology and would like to point out areas that the field needs to advance in. I have shown that SENPs are quite efficient at deconjugating SUMO chains, which they perform in a stochastic manner. This makes SUMO chains quite hard to observe, because they would be constantly cleaved by SENPs in vivo. The schematic diagram depicted in Figure 4.1 displays how SUMO chains could be constantly building up and constantly be deconjugated by SENPs. The mutant that I have utilized, aides in capturing cleavage intermediates and thus has revealed that SUMO chains could be more abundant than previously thought. The fact that they are hard to observe does not mean that they do not exist or have no important signaling properties. Perhaps however, their existence is only crucial under specific circumstances, which still remain out there to discover. Besides the function of polySUMOylation of a protein, several aspects of the proteases that deconjugate them also require further investigation.
Figure 4.1 – Schematic representation of the dynamics & mechanism of SUMO chain deconjugation by SENPs
**The future of SENP research.** Even though it is understood that only a fraction of a given substrate is SUMOylated at a time, suggesting that SENPs are constant "negative" regulators of SUMOylated proteins, the question still remains as to how the regulation of individual SENPs is achieved. Furthermore, their efficiency is highlighted by the now appreciated fact that SUMO chains exist, yet they are nearly impossible to observe experimentally. Determining what the exact biological function of SUMO chains could be will also be a competitive area of SUMO research in my opinion.

The general need for transient SUMOylation/deSUMOylation is in itself a mystery, raising the question of adverse effects resulting from glitches in the pathway, such as altered gene expression due to transcription factor modification or a cellular stress, such as heat shock. Owing to such diligent work done on the subject, the SENP field has progressed at a rapid pace since the discovery of the first *de facto* deSUMOylating enzyme (Li and Hochstrasser, 1999) with great achievements along the way. The genetics from yeast ties in very closely with most mouse knock-out and human cell knock-down experiments that, complemented with classic biochemical analysis of the proteases (Mikolajczyk et al., 2007; Reverter and Lima, 2006), provide a strong basis for studying SENP-mediated deSUMOylating events in cells. State-of-the-art proteomic strategies have also aided in the identification of hundreds of SUMOylation substrates involved in diverse cellular processes leaving many avenues open for future advances (Golebiowski et al., 2009). The proof-of-
principle proteomics techniques presented in Chapter II of my dissertation are excellent complimentary tools to the existing ones in identifying not only (de)SUMOylation substrates, but also the lysine of (de)conjugation as well as identifying novel substrates that are modified by polySUMOylation. Indeed, a recent proteomics study has addressed the profiling of SUMO conjugation sites on a medium-scale (Matic et al., 2010) and has uncovered unconventional SUMOylation consensus sites, indicating the versatility of the SUMO conjugation system. Where the two techniques could synergize is the assignment of specificity regarding deSUMOylation of these consensus sites, which surely would be an important study in the future.

Another future challenge in the field, in my opinion, lies in solving the puzzle of the N-terminal domains of SENPs, shedding light on how these proteases are regulated, and in turn how they regulate global deSUMOylation. Obtaining crystal structures of full-length enzymes will be essential, even though this could turn out to be quite cumbersome to achieve due to the intrinsic flexibility of the N-terminal domains. It is highly possible that the N-terminal domains are unstructured on their own, but would conform to defined tertiary structure when complexed with interacting proteins. Therefore, determining the endogenous interacting partners of full-length human SENPs is another crucial task waiting to be completed. Such putative “SENPosomes” could be key in regulating the activity and specificity of deSUMOylation and/or SUMO chain deconjugation by restricting substrate access or keeping the SENPs in inactive
conformations, perhaps similarly to how the deubiquitinating activity of the SAGA complex is regulated (Kohler et al., 2010).

**SENP**s & **DUB**s & **rock’n’roll.** It would be impossible to speak of SENPs with respect to their activity and mechanisms regarding deSUMOylation and SUMO chain deconjugation and not make a comparison with deubiquitinylating enzymes (DUBs) in cleaving/editing/removing ubiquitin chains.

While currently there is only one *de facto* deSUMOylating enzyme family known, there are nearly one hundred DUBs in humans (Nijman et al., 2005b) encompassing five different protease families (Reyes-Turcu et al., 2009), which is in the ball-park of an impressive one-fifth of all the proteases encoded by the human genome (Lopez-Otin and Overall, 2002). The different families of DUBs have developed various folds and auxiliary domains that direct their specificity towards a particular ubiquitin linkage (Komander et al., 2009) or whether they cleave ubiquitin one at a time or *en bloc* (Reyes-Turcu et al., 2008). Generally speaking, DUBs of the UCH family prefer ubiquitin-conjugates as endopeptidase substrates and prefer to remove C-terminal extension from ubiquitin, whereas members of the USP family usually process larger leaving groups. Such distinctions for the OTU and Josephin family of DUBs are less obvious. These four DUB families are all cysteine proteases with a papain-like fold, while the fifth DUB family, the JAMM proteases are metallo-enzymes. These different DUBs not only prefer various ubiquitin-ubiquitin linkages – which involve all seven lysines of ubiquitin – but some only remove monoubiquitin, such as USP1 from
152

FANCD2 (Nijman et al., 2005a), or they are involved in regulating ubiquitin chain length at the proteasome (Hu et al., 2005). Furthermore, some DUBs have a broader ubiquitin-like molecule specificity and have been shown to cleave not only ubiquitin, but Nedd8 as well, such as USP21 (Gong et al., 2000), or even ISG15, such as USP18 and USP13 (Catic et al., 2007). As more and more mechanistic details were emerging from individual DUBs, it was quite timely to place these proteases into their respective cellular roles through a comprehensive protein-protein interaction study using DUBs as baits, achieved by a massive proteomics study (Sowa et al., 2009). Such a “DUB interaction landscape” provides an excellent starting point for determining the regulation of DUBs by identifying the proteins they interact with, giving clues about the complexes they could be a part of, such as the SAGA complex, which deubiquitinylates histone H2B (Kohler et al., 2010). In principle, such a comprehensive study with seven human SENPs is far less complex than with the plethora of DUBs, therefore it is surprising that we still do not know whether SENPs could be part of large complexes or not. Though, as I mentioned before, this for sure will be the theme of future research.

Another aspect in which our knowledge of SENPs has been lagging behind DUBs is their mechanism of chain deconjugation, whether all SENPs are capable of SUMO chain trimming or whether there is a division of labor among them. However, in the light of my findings in Chapter III regarding the dynamics and mechanism of SUMO chain deconjugation, I should argue that it makes sense for SENPs to have a unified way of recognizing and cleaving SUMO
molecules, whether it is a single SUMO on its own or part of a chain, because they are a small family of proteases all having the same catalytic fold. Attractive and simple an explanation this is, the large N-terminal domains of SENPs – that so far have eluded extensive mechanistic research in terms of their contribution to the activity and specificity of SENPs – could turn out to contain functional SUMO-interacting motifs (SIMs) that would drive their specificity towards polySUMOylated proteins (Kolli et al., 2010). However, as I have shown, this is unlikely the case for SENP1. Currently, only SUMO-targeted ubiquitin ligases (STUBLs) and certain SUMO-binding proteins have been shown to contain SIMs (Perry et al., 2008), but it would be very interesting to see, on the other hand, if certain DUBs contain SIMs as well. Surely, if there are ubiquitin E3 ligases that specifically recognize SUMOylated proteins, then there must be DUBs that do the same, especially considering the possibility that SUMO itself could be tagged by ubiquitin (Matic et al., 2008). If SUMOylated proteins are in fact ubiquitinylated and degraded by the proteasome, then one obvious question is whether SUMO itself is consumed by the proteasome or is also cleaved beforehand and recycled, similar to ubiquitin. Surprisingly, a recent report identified a SUMO-dependent isopeptidase in yeast that recognizes SUMO molecules but could cleave both SUMO and ubiquitin and is possibly associated with the proteasome (Mullen et al., 2010). While this work may require further validation, it opens up a whole new level of crosstalk between ubiquitinylation and SUMOylation, with a focus on the proteases that deconjugate them.
Clearly, the realms of post-translational modifications by ubiquitin-like modifiers are not isolated, but an intricate interplay exists not only at the level of ligation, but also at the removal of UBLs – a sometimes overlooked area of biology that my dissertation allows a small glimpse into, while serving as a testimony that we are just beginning to uncover the details of these fascinating (iso)proteolytic pathways.
REFERENCES


