Discovery and Development of Small Molecule Probes for the SUMO protease SENP1, a Novel Target for Advanced Prostate Cancer Therapy

by

Carrie M. Johnson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2017

Doctoral Committee:

Associate Professor Jorge A. Iñiguez-Lluhí, Chair
Associate Professor Chris E. Canman
Professor Scott D. Larsen
Professor Diane M. Robins
Professor John J. G. Tesmer

Carrie Margaret Johnson

carriemn@umich.edu

ORCID ID: 0000-0002-6922-0875

Acknowledgements

First, I would like to thank my advisor, Dr. Jorge Iñiguez-Lluhí. His mentorship in the classroom and the lab have helped me improve as a scientist and as a person. I came in as a chemist with analytical and synthetic chemistry research experience and very little in research in biology. With his reputation for his quality teaching, I was able to directly benefit from his talents to learn pharmacology, protein biochemistry, molecular cloning, and numerous other skills and topics. I have benefitted tremendously from your teaching and training and your incredible ability to venture into completely new content areas to come up with interesting questions and potential applications to our work. Lastly, thank you for your kindness and patience as I had to find the tenacity to overcome a two-year series of knee injuries that affected more than just my physical abilities to do research. I am not sure if there are adequate words that fully encapsulate my gratitude for your efforts put into me, but I offer my sincerest thank you.

To my thesis committee, for your dedication to helping see me succeed in my project by offering suggestions, and succeed in my career endeavors by intentionally discussing my interests at each committee meeting and individually. I especially want to thank Dr. Scott Larsen and his graduate student Jeffrey Zwicker for their dedicated efforts while collaborating on the project. To Dr. Diane Robins for your mentorship, support, conversation, and availability throughout the highs and lows of my graduate training.

To Dr. Marcelo Murai, the research associate in the lab for the last three years. Thank you for your incredible support in both the lab and in life. I appreciate your

willingness to take time out of your busy schedule to discuss experiments and techniques with me. I will miss your spontaneous little gifts of treats and snacks. You are a critical component of me getting to this point and I am so grateful for your friendship and mentorship.

To my undergraduate mentors, Dr. Dwight Stoll, Dr. Amanda Nienow, Dr. Scott Bur, and Dr. Brenda Kelly, and Dr. Uttam Tambar, I never thought I was smart enough to major in chemistry let alone do research. With your support, you encouraged me to get in the lab and try it out and then later to pursue graduate school. Thank you for believing in me when I never would have. Additionally, thank you for your availability in continuing to serve as mentors beyond my bachelor's degree.

As the child of two teachers in a small community, I have long been aware of the importance of teachers throughout my academic path. From Kindergarten to 12th grade, rural Montevideo provided me with a great education to lay the foundation that allowed me to earn my PhD. I know teachers are not recognized and thanked enough for their dedication and skills in teaching students. Know that I am grateful for your sacrifices to my education.

To my friends of graduate school, ice hockey, FEMMES, and beyond, thank you for your facilitation of incredible memories, your support as a surrogate family throughout my PhD training, and your listening ears and advice through both the cathartic and celebratory conversations. There are too many of you to name, but I gained the best friendships of my life in my time here thanks to you. I truly love this University and what it has fostered. Forever Go Blue!

Finally, I would like to my thank family, including my parents Mark and Dana, my brothers Matthew and Eric, and their wives Julie and Becca, my grandparents, and distant family members Hannah Bolados and Elaine Miller. I am not sure where to really begin, since some of you date back to my beginning, but I know I am so blessed to have such a truly amazing family. Thank you for supporting a culture of activity that provides me with a lifetime of athletic outlets. Thank you for helping me navigate my entire academic career thus far, even if some of you did not always know what this latest step of pursuing a PhD entailed. Thank you for your dedication to my educational pursuits as you sacrificed personal time to talk to me or come visit. This dissertation and degree is your accomplishment as it is mine.

Though there are many others left and this might be one of the longest acknowledgments for a dissertation, I am the product of countless individuals to whom I am grateful. They are too important and I receive great joy in attempting to note them all.

TABLE OF CONTENTS

List of Figures

List of Tables

Abstract

Androgen signaling through the androgen receptor (AR) is essential for normal growth and function of the prostate gland. In prostate cancer (PC), androgens provide the main proliferative drive for the disease, making androgen-deprivation one of the primary therapeutic strategies. Although initially effective, such treatments select for tumor cells that are able to sustain proliferation in a reduced androgen environment. This allows for the emergence of castration resistant PC, an incurable disease where both the AR transcriptional program is subverted and cellular senescence is evaded. SUMOylation is a post-translational modification that regulates both of these processes. SUMOylation of AR inhibits both basal and androgen-stimulated transcription in a promoter context manner and enhanced global SUMO modification induces prostate cell senescence. Advanced PC cells evade these mechanisms at least in part through the upregulation of SENP1, a SUMO-specific cysteine protease that reverses SUMOylation. In addition, AR is a direct activator of the SENP1 gene, creating a self-reinforcing loop that promotes and sustains its own activity and PC progression.

This dissertation is aimed at the discovery and development of small molecule inhibitors of SENP1 as the basis for novel prostate cancer therapeutics. Using a robust FRET-based assay, we defined the kinetic properties of SENP1 and its closest paralog SENP2. This analysis revealed significant product inhibition and a differential sensitivity to ionic strength. Using this assay, an extensive high-throughput screening campaign led to the identification of two structurally distinct inhibitor classes. Characterization of these

compounds indicates that they display significant selectivity towards SENP1 relative to SENP2 and that they act in both a reversible and competitive manner. Furthermore, these compounds inhibit native full-length SENP1 acting on endogenous SUMOylated substrates. Notably, both groups of compounds are known to display activity as purinergic receptor antagonists. The remarkable parallel pharmacology to P2X1 receptors led to our discovery that ATP, the endogenous P2X ligand, is a SENP1 selective inhibitor. We have thus revealed a novel nucleotide mediated regulation of SENP1. Using a combination of mutagenesis, biochemical assays, and fluorescence and NMR spectroscopy, we have characterized the binding of inhibitors and identified key enzyme residues involved in the interactions as well as residues responsible for SENP isoform selectivity.

These findings reveal that SENP1 harbors a unique binding site for nucleotides that can be targeted by small molecules. This knowledge can guide novel strategies for further inhibitor development for evaluation of the therapeutic efficacy of SENP1 inhibitors in advanced prostate cancer models.

Chapter 1

Introduction to post-translational modification by SUMO and its roles in the modulation of function and disease

1. Overview

Despite containing the same copy of DNA, the types of cells in the human body vary widely and execute an even wider array of functions to maintain the homeostasis of our complex biology. The breadth of cell type and function is the product of a variety of mechanisms that maintain exquisite control over the transcription, translation, degradation, and post-translational modifications of protein products. Transcription, the process of generating messenger RNA (mRNA) from parent DNA, and translation, the process of chemically translating mRNA to generate protein, are two of the most fundamental activities of cells. Proteins execute many of the structural, enzymatic, and messenger responsibilities of the cell. In opposition to transcription and translation for protein synthesis is protein degradation. To achieve cellular homeostasis, competition between synthesis and degradation allow cells to balance the levels and types of proteins available in accordance with the needs of the cell. These competing processes are nonetheless essential and vary in response to stimuli. However, their rate of response is not always rapid enough for meeting all cellular needs.

Post-translational modification of proteins adds another layer of control to protein function. The plasticity of these modifications allows for dynamic control of protein

function that can be more rapid than the transcription/translation and degradation options. In turn, transcription and translation themselves can be regulated by post-translational modification of the proteins involved in these processes. There are over 400 unique post-translational modifications (PTMs) that are grouped into 25 identified major categories (Khoury et al., 2011). These PTMs include small chemical groups (such as phosphate, hydroxyl, methyl, and acetyl groups), various lipids, arrays of sugars with many different branching types, and other proteins. Of the 20 natural amino acids, 15 have been observed to be post-translationally modified (Walsh et al., 2005). Altogether, these PTMs can result in a change in the localization, activity, stability, and/or the interactome of the modified protein. Working as an on-off switch or operating within a continuum, most of these PTMs are reversible. This can allow for dramatic and prompt changes in the function of a protein between the two or more states with cell-altering consequences. Due to their significant role in cellular homeostasis and the disease consequences when they go awry, PTMs are of significant interest in biomedical research.

2. Ubiquitin

Among the many types of PTMs and compared to phosphorylation, the discovery and research on protein-based PTMs is still young with many unknowns despite their central role in regulating cell biology functions. Discovered in 1975, the appropriately named ubiquitin is a protein-based PTM that has been found to be conjugated to tens of thousands of proteins (Cappadocia and Lima, 2017; Chen et al., 2014; Schlesinger et al., 1975). With only three conservative mutations from yeast to human, ubiquitin is one of the most invariant proteins across species (Komander and Rape, 2012). Ubiquitin is a small

8.5 kDa protein that modifies other proteins commonly through conjugation of its C-terminal glycine residue to the epsilon-amino group (ε-NH_3^+) of a lysine on the target protein, creating an isopeptide bond. This conjugation is executed through a pathway involving a series of enzymes known as E1, E2, and E3 enzymes (Hershko et al., 1983). E1 enzymes are known as the activating enzymes, where the enzyme activates ubiquitin through an ATP and magnesium-dependent process. Through the action of E1 activating enzymes, ubiquitin is first adenylated yielding AMP at its C-terminus and pyrophosphate. The E1 catalytic cysteine then attacks the C-terminal carbonyl carbon of the adenylated ubiquitin to release AMP and form a thioester intermediate linking E1 and ubiquitin. Lastly, the E1 activating enzymes catalyze the trans-thioesterification of ubiquitin to E2 enzymes. The E2 enzymes, known as the conjugating enzymes, carry the charged ubiquitin and coordinate with an E3 enzyme and a target protein substrate. E2 conjugating enzymes can also directly catalyze the transfer of ubiquitin to target substrate in the absence of an E3 enzyme. The E3 family of ligases contains nearly 600 members and are divided into three recognized classes based on their domains: really interesting new gene (RING), homologous to the E6-AP carboxyl terminus (HECT), and RING in between RING (RBR) (Chymkowitch et al., 2011; Kodadek et al., 2006). The RING class of E3 ligases is the largest and act as scaffolds in coordinating the ubiquitin-linked E2 enzyme and the target substrate for E2-catalyzed transfer. In contrast, the HECT and RBR classes operate under a two-step process where the ubiquitin is passed from the E2 to a catalytic cysteine on the E3, followed by transfer to the target substrate lysine. In the ubiquitin pathway, there are two E1s, around 45 E2s, and over a thousand E3s encoded in the human genome that together help ubiquitinate a multitude of protein targets (Jin et al., 2007; Wenzel et al.,

2011; Zhao et al., 2012). The complexity and number of permutations involved in understanding the specificity of linkage between target proteins to the array enzymes in E1, E2, E3 pathways remains limited.

Embedded in the complexity of ubiquitination pathways and associated targets are the wide range of functional consequences of this modification. In addition to ubiquitination occurring on different unique proteins and on multiple potential sites within some proteins, perhaps ubiquitin's most notable conjugation pattern is ubiquitin-modified ubiquitin. Ubiquitin has seven embedded lysine residues to which an additional ubiquitin can be conjugated, allowing for the formation of ubiquitin chains and different patterns of linkages (Peng et al., 2003; Xu et al., 2009). The importance of chain formation was recognized with the 2004 Nobel Prize in Chemistry, awarded for the discovery of ubiquitin-mediated protein degradation. The most recognized function of ubiquitin is that its chains can lead to protein degradation through the 26S proteasome (Elsasser and Finley, 2005; Saeki et al., 2002). Despite this common understanding, not all chains and not all proteins tagged with ubiquitin are targeted for degradation. The different chain lengths, target proteins, and types of linkages allow for a diversity in topology and therefore function. Like many PTMs, ubiquitin also has roles in regulating many cellular functions, including DNA repair, transcription, endocytosis, cell division, and apoptosis (Huang and D'Andrea, 2006; Hurley et al., 2006; Mukhopadhyay and Dasso, 2007; Sokratous et al., 2014; Vucic et al., 2011).

Deubiquitinating enzymes (DUBs) cleave ubiquitin from substrates and are responsible for generating new free ubiquitin molecules. Through bioinformatics and experimental approaches, nearly 100 DUBs have been identified and they are subdivided

into five classes (Hanpude et al., 2015). Making up four of the five classes, the majority of DUBs are cysteine proteases, with the fifth class being metalloproteases. Four genes, ubiquitin B (UBB), ubiquitin C (UBC), ubiquitin A-52 residue ribosomal protein fusion product 1 (UBA52), and ubiquitin-40S ribosomal protein S27a (UBA80), serve as the source for generating new free ubiquitin. These genes are transcribed and translated as linear tandem ubiquitin fusions or as N-terminal fusions to other proteins. To become amenable for serving as a PTM, these ubiquitin fusion molecules are processed through DUB cleavage for cytosolic release. Ubiquitin is recycled through the actions of a DUB. Integrated into the 26S proteasome that cleaves ubiquitin from proteins targeted for degradation. In the non-proteasomal setting, the reversible nature of ubiquitination becomes essential in tuning the activity, localization, and other functions of ubiquitin-modified proteins. DUBs can modify ubiquitin linear and branched chains that generate unique signals, thereby changing those signals and even rescuing proteins from degradation (Komander et al., 2009; Schrader et al., 2009). While much research remains to be completed to achieve a more thorough understanding of the regulation of DUBs, it is clear that PTMs play an important role in the localization, stability, and activity of these enzymes. Phosphorylation, ubiquitination, redox modification, and SUMOylation (a PTM to be addressed shortly) of DUB(s) have been observed and the functional consequences are beginning to emerge (Hanpude et al., 2015). For example, phosphorylation of the DUB ataxin-3 occurs in one of its ubiquitin-interacting motifs (Mueller et al., 2009). Spinocerebellar ataxia type 3 (SCA3) is a neurodegenerative disease caused mutations in ataxin-3 which yield an expanded polyglutamine track in the protein and is associated with the accumulation of neuronal intranuclear inclusion bodies. Phosphorylation of pathogenic

ataxin-3 controls ataxin-3's stability, aggregation, and localization in a way that enhances its nuclear localization and formation of nuclear inclusions (Mueller et al., 2009). Ubiquitination of ataxin-3 at a specific lysine, K117, enhances ataxin-3's DUB activity (Zhou et al., 2013). In a third type of PTM that alters the disease status, ataxin-3's stability and cytotoxicity is enhanced by a process called SUMOylation.

3. SUMO

Following the discovery of ubiquitin, other ubiquitin-like proteins (Ubls) were soon discovered. They are designated as type I or type II based on whether they are conjugated to protein substrates (I) or not (II). These Ubls are typically small like ubiquitin (~10 kDa) and share a common structural fold consisting of a β-grasp domain of five antiparallel β-strands partially wrapping an α-helix within a ββαββαβ secondary structure sequence (Bayer et al., 1998; van der Veen and Ploegh, 2012; Xu et al., 2009). Today, there are a total of 17 type I Ubl proteins, broken down into eight known classes: SUMO, NEDD8, ATG8, ATG12, URM1, UFM1, FAT10, AND ISG15 (Cappadocia and Lima, 2017).

Just over two decades ago, a Ubl was discovered and independently named by four separate labs. Ultimately the name Small Ubiquitin-like Modifier (SUMO) has been commonly adopted (Boddy et al., 1996; Matunis et al., 1996; Okura et al., 1996; Shen et al., 1996). Among the eight classes of type I Ubls, SUMO is the most pervasive and is conserved from yeast to mammalian cells (Chen et al., 1998). SUMO and ubiquitin share 48% structural fold similarity and 18% sequence homology (Figure 1-1A) (Bayer et al., 1998; Wilson, 2009). Though they share a similar fold, the surface of SUMO that contacts some of the essential enzymes in the SUMO pathway is more negatively charged than

ubiquitin (Figure 1-1B) and helps differentiate among the species that recognize the similar fold. In the SUMO class of Ubls, there are four mammalian SUMO variants on separate genes, SUMO1, 2, 3, and 4 (Bohren et al., 2004; Kamitani et al., 1998). SUMO1 shares 46% identity and 66% homology with SUMO3 while SUMO2 and SUMO3 are nearly

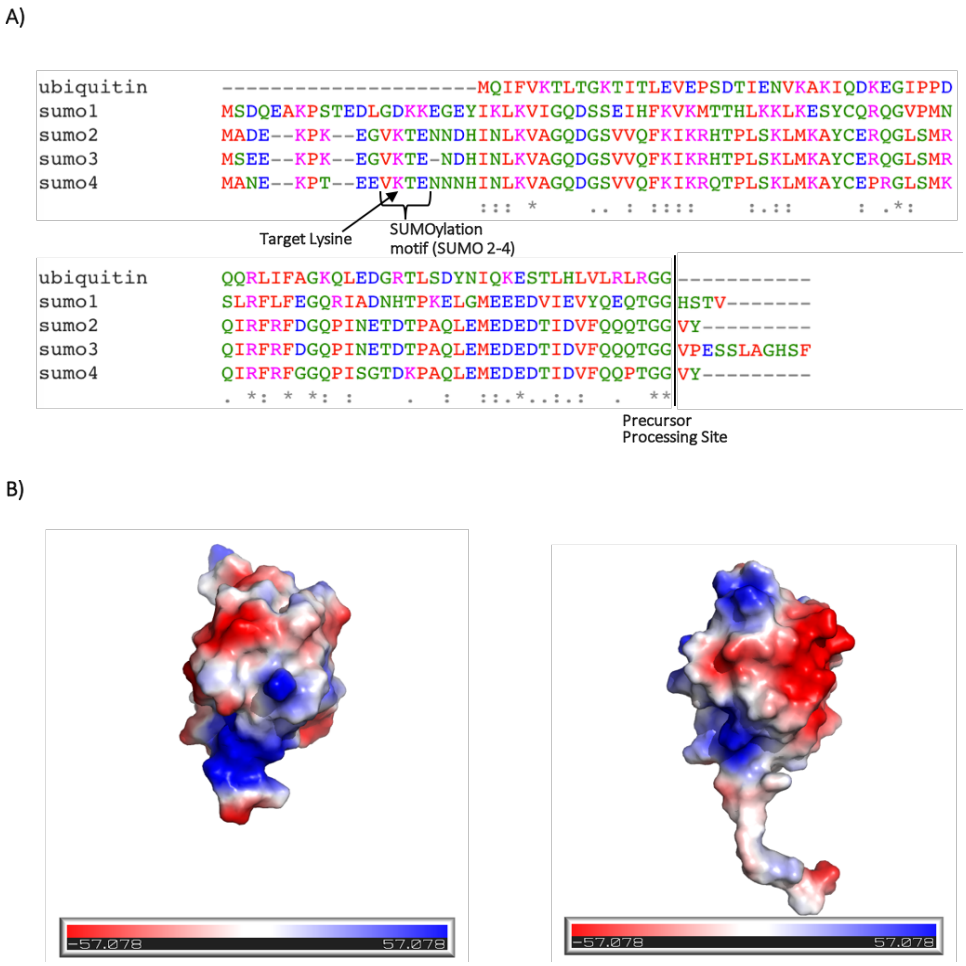


Figure 1-1 Sequence and structural surface comparison of ubiquitin and SUMO.
A) Sequence alignment showing the similarity between ubiquitin and the four SUMOs. Basic residues shown in magenta, acidic in blue, small hydrophobic (phenylalanine and tryptophan) in red, and the remainder (hydroxyl, sulfhydryl, amine, and glycine) in green. The lysine residue noted in SUMO2-4 is the residue SUMOylated in SUMO chains. Sequence Alignment completed using Clustal Omega: Larkin et al. B) Electrostatic potential of ubiquitin (left) and SUMO1 (right) showing the face that interacts with their respective proteases (USP2 and SENP1 respectively). PDB ID: 3V6E (left) 2IY1 (right). In terms of vacuum electrostatics, red coloring indicates negative potential and blue indicates positive potential. This negative-positive coloring scheme correlates with the positive attributes of the University of Michigan (Go Blue!!) relative to its rival Ohio State.

identical with 92% identity and an absence of antibodies capable of distinguishing the two (Kamitani et al., 1998). As a result, they are often referred to as SUMO2/3. Notably, like ubiquitin, SUMO2/3 contain their own internal lysine residue capable of being modified by SUMO and thereby allowing for the formation of SUMO chains *in vivo* (noted in Figure 1-1A (Tatham et al., 2001). SUMO1, 2, and 3 are expressed in all tissues that have been tested (Wilson, 2017). In contrast, SUMO4, which shares 87% homology with SUMO2, is expressed primarily in the kidney, spleen, and lymph (Bohren et al., 2004; Guo et al., 2004). Within the SUMO family, only SUMO2 is essential. Despite the similar identity between SUMO2 and SUMO3, SUMO3-deficient mice are viable whereas SUMO2-deficient mice die at embryonic day 10.5 (Wang et al., 2014). The difference in viability does not appear to be due to functional differences between the two SUMOs but rather a result of the difference in their unique expression pattern during embryonic development. Out of the three prominent SUMOs, SUMO2 makes up 75-80% of mRNA levels and SUMO3 only makes up 2-3% during embryonic development (Wang et al., 2014). SUMO1-deficient mice are viable and SUMO2/3 can compensate for loss of SUMO1 (Evdokimov et al., 2008; Zhang et al., 2008).

Like ubiquitin, SUMOs are translated as a precursor that must be proteolytically processed to expose the flexible C-terminal di-glycine motif (Figure 1-1A). This is the mature form that is amenable to conjugation to substrate proteins. This precursor processing is performed by a set of proteases called SENPs, which will be further described shortly. Once processed to mature SUMO, conjugation of SUMO to substrate, termed SUMOylation, takes place via a set of dedicated E1, E2, and E3 enzymes that are similar in enzymology to those in the ubiquitin system (Figure 1-2). There is one known E1

activating enzyme complex for the SUMO system, a heterodimer of SUMO activating enzymes 1 (SAE1) and 2 (SAE2) (Desterro et al., 1999; Gong et al., 1999). In contrast to the ubiquitin system, there is only one known SUMO E2 conjugating enzyme, termed ubiquitin conjugating enzyme 9 (Ubc9) or ubiquitin conjugating enzyme E2I (Ube2I) (Johnson and Blobel, 1997; Watanabe et al., 1996). Ubc9 is essential for development, as Ubc9-deficient mice are embryonic lethal, dying early after implantation (Nacerddine et al., 2005). There are a handful of E3 ligase enzymes for SUMO, but unlike the

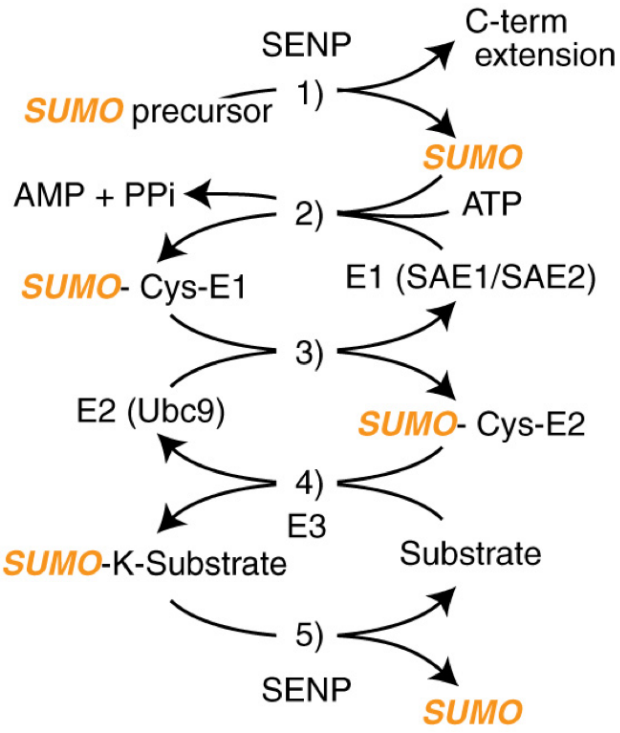


Figure 1-2 SUMOylation pathway.
The SUMO precursor protein is processed by a member of the SENP enzyme family (1) to generate a SUMO capable of adenylation by the SUMO E1 enzyme complex (SAE1/SAE2) and conjugation to its catalytic cysteine (2). The SUMO is handed off to the SUMO E2 enzyme Ubc9 via transthioesterification to its cysteine residue and the E1 enzyme can cycle back to activate new SUMO molecules (3). With the help of a SUMO E3 enzyme, SUMO is transferred from Ubc9 to its target substrate lysine (4) embedded within a SUMOylation motif. Reversal of SUMOylation is carried out by the SENP enzymes, which deconjugate (deSUMOylate) it from the target protein (5) and recycle the free SUMO to be used for subsequent rounds of SUMOylation.

ubiquitination pathway, they are not always necessary. SUMOylation of target proteins can be achieved *in vitro* without the aid of an E3 ligase (Desterro et al., 1999; Okuma et al., 1999). SUMO E3 ligases, like protein inhibitor of activated STAT1 (PIAS1) and Ran binding protein 2 (RanBP2), generally enhance SUMOylation efficiency and specificity by bringing the SUMO-loaded Ubc9 and the SUMO-targeted substrate together or by aligning the SUMO for efficient conjugation without necessarily interacting with the substrate itself (Gareau and Lima, 2010; Kahyo et al., 2001; Pichler et al., 2002). As a mechanism to bypass the E3 ligases, Ubc9 contains a surface near its active site cysteine that is able to bind the SUMOylation consensus motif in an independent manner (Bernier-Villamor et al., 2002).

The consensus sequence on substrate proteins for SUMOylation is ψKxD/E and is often flanked by proline or glycine residues, suggesting a loop or unstructured region accompanies the SUMOylation motif (Iñiguez-Lluhí and Pearce, 2000; Mukherjee et al., 2012; Sampson et al., 2001). In this consensus sequence, ψ represents a hydrophobic residue, K is the lysine containing the ε-NH_3^+ group for isopeptide linkage to SUMO, x is any amino acid, and the motif ends with an acidic residue. Despite this canonical motif, numerous proteomics studies have yielded alternative motifs as SUMOylation sites. Among these alternatives are an inverted SUMO motif (D/ExKψ), which notably is found in SUMO1, an expansion of the preceding hydrophobic residues (ψψψKxD/E), and motifs that include SUMOylation sites dependent on downstream negative charges through phosphorylated serines or acidic amino acids (Hietakangas et al., 2006; Matic et al., 2010; Yang et al., 2006). With these sequences in mind, the best proteomics data puts the number of observed SUMOylated human proteins at 3,617 with 7,327 total SUMOylation sites and

a maximum of 63 sites on one protein, E3 SUMO-protein ligase zinc-finger 451 (ZNF451) (Hendriks and Vertegaal, 2016). With its E3 ligase activity and E4 elongase activity, meaning it elongates SUMO chains, ZNF451 is thought to experience self-SUMOylation with the aid of its multiple SUMO-interacting motifs (SIMs) (Eisenhardt et al., 2015; Hendriks and Vertegaal, 2016). The E2 conjugating enzyme Ubc9 has its own SUMO consensus motif that is essential to its function in modifying substrates with SUMO (Sampson et al., 2001).

Like other PTMs, SUMO is reversible. DeSUMOylation, or cleavage of the isopeptide bond between the substrate ε-NH_3^+ group and the carboxyl group of mature SUMO's terminal glycine, is performed by the previously mentioned SENP enzymes. With their name rooted in one of SUMOs original names, sentrin, these sentrin or SUMO proteases are a family of cysteine proteases that catalyze both the SUMO precursor processing and the removal of SUMO from its substrate. There are six SENP enzymes, SENP1-3 and SENP5-7. They share a common C-terminal catalytic domain containing the His and Asp residues as part of the catalytic triad essential for cysteine-mediated peptide hydrolysis but contain variable N-terminal domains (NTDs) that regulate their localization and specificity (Gong et al., 2000; Hang and Dasso, 2002; Kim et al., 2000; Nishida et al., 2000, 2001). The differences in localization, precursor processing (or endopeptidase activity), deconjugation (or isopeptidase activity), chain-editing (protease activity on SUMO chains) and SUMO paralog specificity are outlined in Table 1.2.1 (Hickey et al., 2012; Kolli et al., 2010; Mendes et al., 2016; Mikolajczyk et al., 2007; Nayak and Müller, 2014).

Table 1.2: SENP characteristics.

	Localization	SUMO preference	Precursor Processing	deSUMOylation	Chain editing
SENP1	Nuclear pore, nuclear foci, and cytoplasm	SUMO1-3	Yes	Yes	Yes
SENP2	Nuclear pore, nuclear foci, and cytoplasm	SUMO2/3 >> SUMO1	Yes	Yes	Yes
SENP3	Nucleolus	SUMO2/3	No	Yes	No
SENP5	Nucleolus & Mitochondria	SUMO2/3	No	Yes	No
SENP6	Nucleoplasm	SUMO2/3	No	Yes	Yes
SENP7	Nucleoplasm	SUMO2/3	No	Yes	Yes

Based on their characteristics, the SENP enzymes are further subdivided in to three groupings. SENP1 and SENP2 have the broadest substrate specificity in both precursor processing and isopeptidase activity across the SUMO family. They also contain a nuclear localization signal within their sequence, and SENP2 contains a nuclear export signal, allowing them to shuttle between the nucleus and cytoplasm (Bailey and O'Hare, 2004; Itahana et al., 2006). Making up the second subclass, SENP3 and SENP5 are nucleolar and prefer SUMO2/3 substrates. Lastly SENP6 and SENP7 form a group and have an extra conserved sequence insertions in their catalytic domains (Nayak and Müller, 2014). Both SENP1 and SENP2 are essential, as mice deficient in either are embryonic lethal (Cheng et al., 2007; Kang et al., 2010; Sharma et al., 2013). In the management of SUMO1-

modified proteins in mice, SENP1 is essential for deSUMOylation, including deSUMOylation of SUMO1-capped SUMO2/3 chains (Sharma et al., 2013). While SENP1 is not essential for deSUMOylating SUMO2/3, as the other SENP enzymes can compensate, SENP1 itself is capable of being modified by SUMO2/3, though the functional consequences of this SUMOylation of SENP1 are unknown (Sharma et al., 2013).

Unlike SUMO1-3, SUMO4 is not processed by an SENP enzyme due to a proline residue found two residues upstream of the C-terminal Gly-Gly motif and is only processed by a stress induced hydrolase (Owerbach et al., 2005; Wei et al., 2008). This lack of precursor processing because of SUMO4's proline is, in part, likely due to the strict steric constraints of SENPs. The catalytic cysteine in SENP enzymes sits under a narrow tunnel created by tryptophan residues (Figure 1-3). This catalytic tunnel creates rigid specificity where only the small glycine residues are able to fit and the proline-induced limitation on secondary structure and phi and psi angles may prevent proper orientation. In conjunction with the E1, E2, and E3 enzymes in the SUMO pathway, the SENP enzymes' functions yield a framework for SUMO recycling and a PTM system with the plasticity that allows for dynamic regulation of cellular processes (Figure 1-2).

4. SUMO Functions

As a macromolecule, SUMO protein can exert its functions through a variety of ways. Via SUMOylation, protein interactions with other proteins can become blocked by occlusion of the surfaces through which they form a complex. In contrast, the new hybrid surfaces created on a SUMOylated protein can promote new protein-protein interactions.

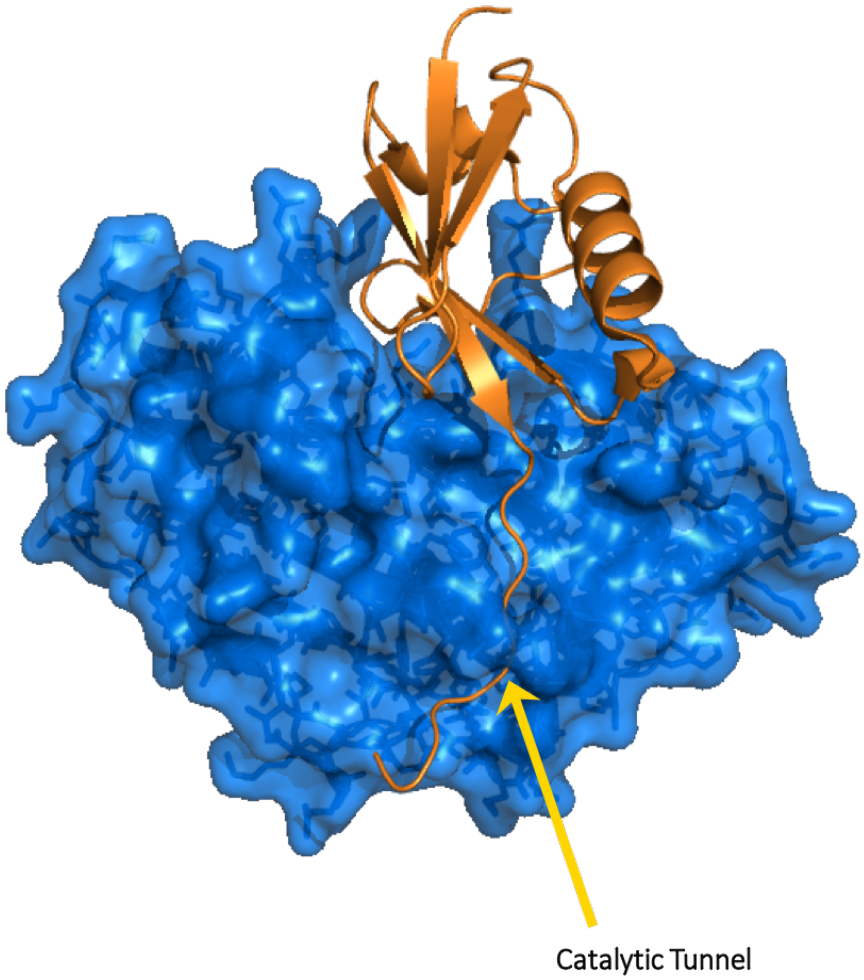


Figure 1-3 SUMO – SENP1 complex.
Crystal structure of SENP1 (blue surface) in complex with SUMO1 (orange ribbon). The catalytic region is indicated by the arrow. The catalytic cysteine is covered by the tryptophan tunnel. The narrow tunnel is able to accommodate the size of the small glycine amino acids. PDB ID: 2IY1.

Additionally, SUMOylation can perturb the tertiary structure of the protein it is modifying, thereby altering the conformation and exposing new interaction surfaces. Lastly, SUMO functions can prevent modification of a SUMOylated lysine by other PTMs. Nearly a third of identified SUMOylation sites overlap with known methylation, acetylation, and

ubiquitination sites, with ubiquitin making up the majority of the sites (Hendriks and Vertegaal, 2016).

To elaborate on one of the mechanisms listed above through which SUMO executes its function, protein-protein interactions introduced via SUMOylation are often facilitated by SUMO interaction motifs (SIMs). Though these protein-protein interactions through SIMs are weak, ~1-100 μM, they have been found on many E3 ligases and other proteins, sometimes with numerous SIMs being found on one protein (Hecker et al., 2006; Keusekotten et al., 2014). Found on the surface of interacting proteins, SIMs are made up of a series of three to four hydrophobic residues and are typically followed by a series of acidic amino acid residues that vary in length (Hannich et al., 2005; Hecker et al., 2006; Kaur et al., 2017; Minty et al., 2000; Song et al., 2004). Together this string of residues forms a β-strand that forms an intermolecular beta-sheet with the β_2 strand of SUMO and interlaces between this β_2 strand and the α_1-helix of SUMO. Notably, this region of SUMO is one of the regions where the sequences of the SUMO isoforms differ substantially and could drive some of the selectivity they exhibit between SUMO isoforms and their functional effects (Huang et al., 2004). Within the series of hydrophobic residues, there is often an acidic residue interrupting it. Additionally, serine and threonine residues are frequently found within the series of acidic residues and/or within the hydrophobic series. These additional features help further guide the selectivity and orientation of the SIM-containing protein to SUMO paralogs (Chang et al., 2011; Hecker et al., 2006; Kaur et al., 2017; Namanja et al., 2012; Song et al., 2005). The acidic amino acid series helps orient the SUMO-SIM interaction, dictating whether the β-strand of the SIM runs parallel or antiparallel to the β_2-strand in SUMO. The serine and threonine residues are potential

phosphorylation sites and are important for preferential binding of SUMO1 over SUMO2. The non-conserved N-terminal domains (NTDs) of SENPs contain one or more SIMs to possibly recruit SUMO, but the SUMO binding surface that SENPs ultimately engage with for catalysis is distinct from the SIMs and covers a larger surface area than the E1 and E2 enzymes (Hickey et al., 2012; Kerscher, 2007; Komander et al., 2009; Namanja et al., 2012). Counter to the SIM surface is the specific surface on SUMO that interacts with the SIMs. Previously, our lab systematically examined the residues on the SUMO2 surface for functional inhibition of transcription (Chupreta et al., 2005). Among the seven residues on SUMO identified in altering SUMO-mediated inhibition, four were positively charged. These seven residues are found on the second beta-strand and the alpha helix following it and are in agreement with the SIM-focused work described above. The location of the surface on SUMO2 that interacts with SIMs is opposite (180°) to the face that interacts with Ubc9 and SENP enzymes.

As might be expected for a PTM that potentially modifies nearly a fifth of the human proteome, the roles of SUMO in cellular function are extensive. Unlike most other PTMs, SUMOylation occurs predominantly in the nucleus where the SUMOylation machinery is typically found (Rodriguez et al., 2001; Zhang et al., 2002). Although SUMO's name is rooted in its similarity to ubiquitin, SUMOylation has wider-ranging functions than ubiquitin. In a macroscopic perspective, SUMO proteomics studies have revealed that there are functional clusters of SUMOylated proteins. These clusters include proteins involved in the DNA damage response, pre-mRNA splicing, ribosomal biogenesis, cell cycle regulation, transcriptional regulation, and chromatin remodeling (Hendriks and Vertegaal, 2016). Because of the promiscuity and lack of diversity in

SUMO E1, E2 and E3 enzymes with respect to the broad functional consequences of SUMOylation, the specificity of SUMOylation targets and functional outputs is thought to be driven by the clustering of proteins collectively responsible for specific categories of tasks (Hendriks and Vertegaal, 2016; Psakhye and Jentsch, 2012). Essentially, with only one identified E1 and E2 enzyme, specificity is supported by locally concentrated substrates for group modification.

Working down from proteomics to a more specific functional perspective, SUMO is intimately involved in DNA damage response pathways. DNA damage in terms of G-U/T mismatch or methylated cytosine is repaired in part by thymidine DNA glycosylase (TDG). After performing the first step of generating an abasic site in the DNA via hydrolysis of the N-glycosidic bond between the base and the deoxyribose, TDG must dissociate from the DNA. This dissociation is the rate-limiting step of its function. As an example of SUMOylation-induced conformational change, after base excision, SUMOylation of TDG generates a conformational change. This SUMOylation reduces the severe product inhibition on TDG by abasic sites and induces dissociation of its glycosylation domain from the DNA backbone while retaining active site integrity and glycosylation function (Steinacher and Schär, 2005; Waters et al., 1999). This aided dissociation creates a significant increase in enzymatic turnover and enhances the effects of enzymes involved in the next steps of base-excision repair (Hardeland et al., 2002; Smet-Nocca et al., 2011).

When double strand DNA breaks arise and repair occurs via homologous recombination (HR), SUMO is again involved, this time modifying a large cluster of proteins. The process of HR is multi-stepped and involves the initial identification and

stabilization of the ends of the free DNA, formation of single-stranded DNA by nucleases, capping of ssDNA, homology searching, and recombination and annealing to complete the repair. Throughout this process numerous proteins are involved, including replication protein A (RPA), Rad52, and Rad59, and the majority of them are SUMOylated (Cremona et al., 2012; Dou et al., 2010; Psakhye and Jentsch, 2012). In the same group-style modification mentioned earlier, DNA damage induces a wave of SUMOylation that enriches with the proteins involved in HR (Psakhye and Jentsch, 2012). Supported by the observance of little to no effect of involved single-protein SUMOylation mutants on HR, the comprehensive SUMOylation profile on the family of proteins involved in HR is key, allowing for synergy among the SUMO modifications (Ohuchi et al., 2008; Psakhye and Jentsch, 2012).

In addition to the effects SUMO has in mediating the activity and function of proteins directly involved in HR, SUMO also plays a role in recruiting ubiquitin E3 ligases for regulation of turnover of proteins involved in HR. SUMO-targeted ubiquitin ligases (STUbL) are a set of ubiquitin E3 ligases that recognize SUMO and through their interaction with SUMO, ubiquitinate the target protein. RNF4 is a STUbL that has four SIMs that can recognize both a series of individual SUMOylations or SUMO chains (Keusekotten et al., 2014). In response to DNA double strand breaks in humans, RNF4 is recruited through its SIMs to ubiquitinate proteins involved in the repair cascade, including RPA and MDC1, and target them for proteasomal degradation (Galanty et al., 2012; Vyas et al., 2013). Additionally, the SUMO-ubiquitin chains created by RNF4 allow for further recruitment of SIM-containing proteins in the BRCA1-A complex which is involved in following steps of DNA damage repair (Guzzo et al., 2012).

In direct opposition to the action of RNF4, ubiquitin specific protease 11 (Usp11) is a DUB that interacts with RNF4 and is targeted specifically to hybrid ubiquitin-SUMO chains (Hendriks et al., 2015). In undoing the action of RNF4, Usp11 removes the ubiquitin polymers from SUMO chains, thereby forming a mechanism to balance the modification and resulting function of the modified protein. In another example of a SUMO deubiquitinase, Usp7 is a DUB present in the chromatin around replisomes. Mature chromatin is ubiquitin rich and SUMO poor, but in the replisome the inverse occurs via the action of Usp7 on ubiquitin modified SUMO2 chains (Lecona et al., 2016).

While SUMO and ubiquitin can work in concert as described above, they can also compete with each other for lysine residues on target proteins. Among the best-characterized examples is the competition between ubiquitination and SUMOylation on Lys21 of the protein IκBα. As an inhibitor of NFκB activity, IκBα complexes with NFκB in the cytoplasm and activation of NFκB involves ubiquitination and subsequent degradation of IκBα. When IκBα is SUMOylated, it is unable to be ubiquitinated and allows for continued repression of NFκB-dependent transcription (Desterro et al., 1998). As an additional layer of cross-talk among PTMs, IκBα phosphorylation is required for its ubiquitination whereas this prevents its SUMOylation. Despite elucidation of SUMO cross talk and competition with other PTMs and point mutants generating further understanding, there remains much to be understood regarding the complex functional consequences surrounding SUMO.

5. SUMO and Transcription

Among the many cellular functions in which SUMO interjects is in transcriptional regulation. It is also among the earliest studied functions of SUMO and has a positive reputable history in our laboratory. With transcription occurring in the nucleus, it is worth noting that nearly 80% of proteins known to be a part of nuclear complexes are SUMO targets (Hendriks et al., 2017). Regulation of transcription by SUMO includes effects on both transcription factors (TFs) and co-regulators in the nucleus. Additionally, this regulation extends beyond the nucleus, interfering with processes coordinating transcription factor activation or degradation in the cytosol. As a result of the essential nature of transcription and effects of SUMO in manipulating it, there have been numerous efforts to elucidate the details of SUMO's role in transcription.

In the assembly of the pre-initiation complex for transcription, co-regulators are proteins that do not directly bind to the DNA and instead bind to and link the transcription factors and RNA polymerase to tune transcriptional output. Co-regulators can act as either co-activators or co-repressors and can be modified by SUMO to disrupt their function or recruit other co-regulators. Many co-regulators have been identified as SUMO targets, including nuclear receptor co-repressor (N-CoR), p300 co-activator, peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α) and receptor-interacting protein 140 (RIP140) (Girdwood et al., 2003; Rytinki and Palvimo, 2008, 2009; Tiefenbach et al., 2006; Treuter and Venteclef, 2011). Depending on the co-regulator and its assigned function, SUMOylation can enhance a co-repressor's repression, de-repress a co-repressor to stop its function, repress a co-activator or prevent its binding, and still in others it can recruit other co-regulators that repress transcriptional activity. In the case of p300, it

typically helps activate transcription through its histone acetyltransferase (HAT) activity, but SUMOylation of p300 antagonizes this function. SUMOylation of p300 recruits histone deacetylase 6 (HDAC6) which reverses acetylation by p300. This deacetylation condenses DNA nucleosomes and represses transcription by p300 (Girdwood et al., 2003).

In addition to co-regulators, TFs are also SUMO targets. Recent Gene Ontology analysis of SUMO proteomics suggests that nearly 300 DNA-binding TFs are SUMO targets, representing over 50% of all TFs (Hendriks et al., 2017; Rosonina et al., 2017). The general understanding of the field is that SUMOylation of TFs reduces transcription of their target genes. Most examples of SUMOylation of TFs yield an inhibitory effect on transcription. Yet there are a few exceptions to this reduction as well as examples of contradictory observations that cloud the understanding of the effects of SUMO on the same TF (Berta et al., 2007; Cheng et al., 2007; Chymkowitch et al., 2015; Gill, 2005; Lyst and Stancheva, 2007). To use the transcriptional co-activator p300 in an example again, SUMOylation of the TF Bcl11b fosters its interaction with p300. In this case, SUMOylation and the presence of p300 causes HAT-driven decondensation of chromatin, relieving the repression of Bcl11b target genes (Zhang et al., 2012). As another example of SUMO-mediated activation of transcription (but without implication of a co-regulator), activation of the Hedgehog pathway results in SUMOylation of the TF Gli1 (Cox et al., 2010). Under conditions where the Hedgehog pathway is not activated, Gli TFs are phosphorylated and subsequently ubiquitinated for proteasomal degradation (Jiang, 2006). Gli1 SUMOylation limits its phosphorylation and ubiquitination (including overlapping targeted lysine residues between SUMO and ubiquitin), lifting the restraints on Gli activity and essentially enhancing relative Gli1-mediated transcription. Regarding contradictory

observations, in some cases SUMOylation of the TF hypoxia inducible factor-1α (HIF1α) has reduced its transcriptional activity yet in others it reduces its activity and leads to its degradation (Berta et al., 2007; Cheng et al., 2007). Hypotheses concerning these types of poorly understood SUMO-effects suggest that there may be tissue specific or other condition specific components tuning SUMO function on TF activity. Nevertheless, the profound impact SUMO has on transcription provides the continued motivation to study and understand the intricacies despite confounding results.

6. SUMO and Nuclear Receptors

Nuclear receptors (NRs) are a superfamily of transcription factors that bind sequence specific promoter elements on target genes. The 48 known NRs in the human genome are subdivided into six evolutionary groups, and half of all NRs have known ligands (Germain et al., 2006). All NRs consist of a modular structure of five to six domains labeled A through F (Figure 1-4A). Working from the C-terminus, the E/F domain consists of the ligand-binding domain (LBD). The LBD mediates ligand binding, receptor dimerization on the DNA response elements, and ligand-dependent transactivation through its activation function 2 (AF-2) helix. Upon ligand binding, the AF-2 helix surface becomes accessible for co-regulator recruitment and binding (Figure 1-4B). The D domain is a poorly conserved hinge region between the LBD and the C-domain and it can contain a nuclear localization signal (NLS). The C-domain, or DNA-binding domain (DBD) is found in all but two NRs and is responsible for binding to specific DNA sequences, termed hormone response elements (HREs). Composed of two cysteine-rich zinc finger motifs and two alpha helixes, the DBD is highly conserved and also participates in receptor

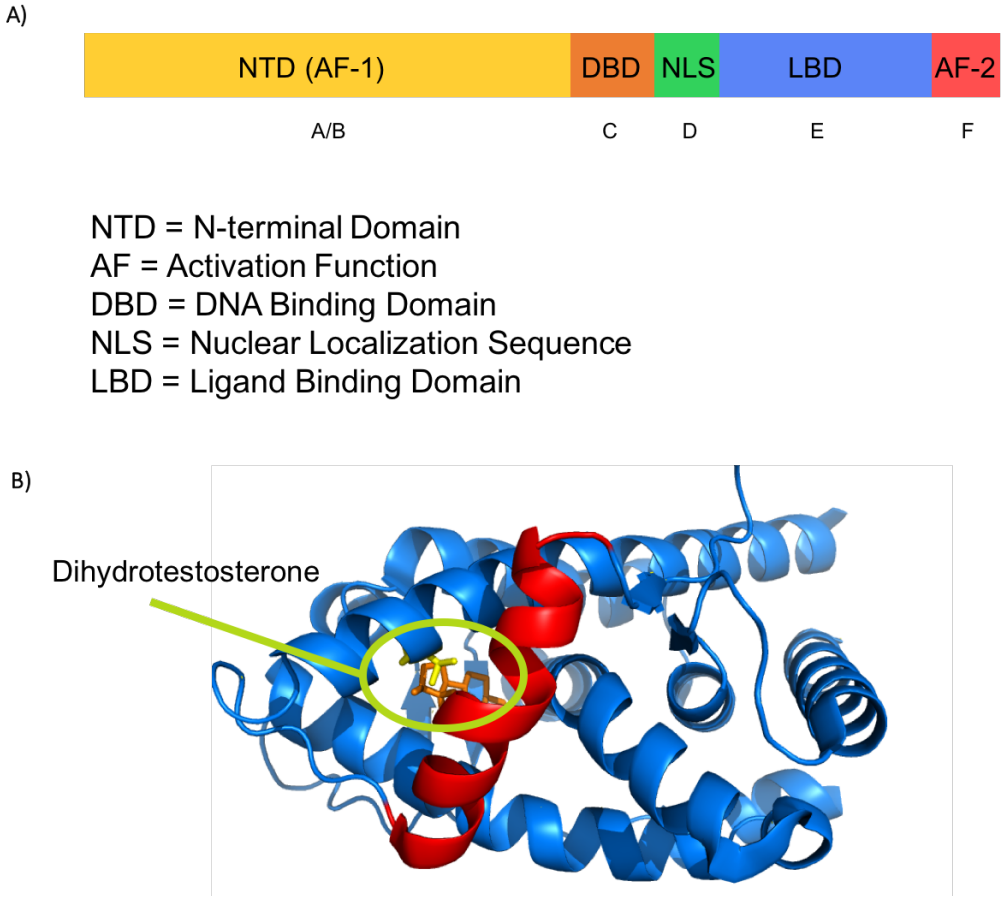


Figure 1-4 Nuclear receptor organization.
A) General domain organization for nuclear receptors from amino to carboxyl-terminus. AF-1 and AF-2 are short for activation function 1 and 2, with the other abbreviations noted. AF-1 is part of the larger A/B domain, with the other individual domains color-coded and labeled. B) Crystal structure of the ligand-binding domain (LBD) of the androgen receptor bound to one of its agonists, dihydrotestosterone (DHT). Helix 12 (red) is essential for the AF-2 function.

dimerization (Germain et al., 2006). The N-terminal A/B region is poorly conserved across the superfamily and contains a transcriptional activation function (AF-1). AF-1 is able to function autonomously, allowing for constitutive activity even in ligand-dependent NRs. The highly variable A/B domain is the target of many PTMs but its structure in any NR has yet to be elucidated.

Like SUMO, NRs play an essential role in nearly all aspects of mammalian physiology and their regulation is necessary to maintain homeostasis. Modification of NRs by SUMO to alter their localization, stabilization, DNA binding, and/or function is a mechanism for regulation of NRs and ultimately their transcriptional output. Peroxisome proliferator-activated receptor (PPAR) α is a NR that modulates metabolism regarding lipid homeostasis and has its transcriptional activity enhanced by binding of co-activators such as p300 or steroid receptor coactivator 1 (SRC1). SUMOylation of PPARα generates a new surface for docking with N-CoR to repress target gene transcription (Pourcet et al., 2010).

As one of the six NR groups, the steroid receptor family of NRs includes the glucocorticoid receptor, estrogen receptors, and androgen receptor. It was in the glucocorticoid receptor (GR) that synergy control motifs, which ultimately became known as SUMOylation motifs, were discovered by our group. In the A/B N-terminal domain (NTD) of GR, a motif within a negative regulatory region was identified to disrupt the transcriptional synergy that occurs when GRs bind to their multiple HREs at a gene promoter (Iñiguez-Lluhí and Pearce, 2000). Compared to a single HRE where one NR dimer can bind, these multiple HREs, termed compound response elements, allow for recruitment of multiple NR dimers. The transcriptional output in response to multiple steroid receptors binding is more than just the additive effect of a single steroid receptor HRE binding, but instead amplifies the output, due to synergy between the individual receptors. Synergy control (SC) motifs limit the scope of this synergy, allowing for tuning of the magnitude of transcriptional response. Mutation of these motifs disrupts the repressed synergy without disrupting DNA binding, but only at compound HREs. These

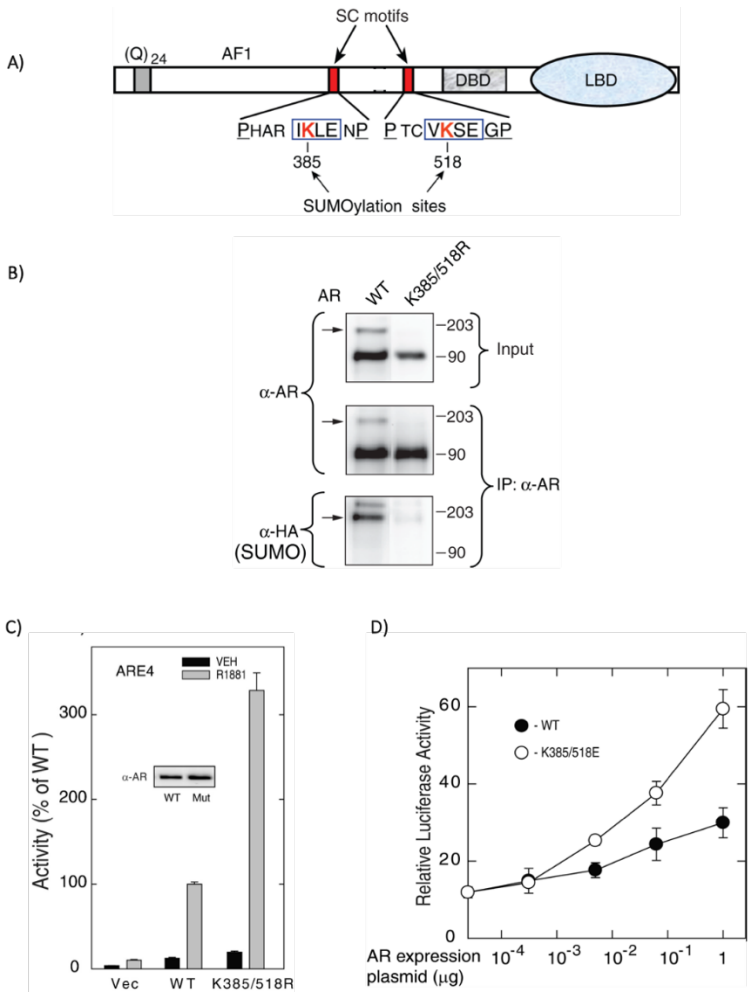


Figure 1-5 AR SUMOylation and transcriptional consequences.
A) Domain organization of androgen receptor (AR) with the synergy control (SC) motifs and the SUMO-targeted lysine within them in red. AF-1 = Activation Function 1, (Q)24 designates the location of the polyglutamine tracks in AR, DBD = DNA binding domain, LBD= ligand binding domain. B) Immunoblot of AR immunoprecipitates from 293T cell extracts transfected with HA-tagged SUMO3 and with either WT AR or AR with a double lysine mutant in the SC motifs (K385/518R). Cells were treated with agonist for 24 hours before lysis. Blots for SUMO or AR are designated and the SUMOylated AR species designated with arrows. C) Normalized reporter assay-based transcriptional activity at compound AREs (4 ARE binding sites in this example) from cells transfected with vector plasmid, WT AR, or AR double lysine mutant in the presence of synthetic androgen (R1881) or vehicle. Inset to show expression levels of WT and mutant AR. D) Luciferase reporter assay based transcriptional activity of WT and double lysine mutant AR in the absence of agonist as a function of the amount of plasmid transfected. Adapted in part from Mukherjee *et al*. 2012.

SC motifs were found in other steroid receptor NRs in their NTD and in other TFs, with different numbers of copies in each (Danciu et al., 2012; Iñiguez-Lluhí and Pearce, 2000).

The androgen receptor (AR) is one of those steroid receptor NRs with SC motifs.

The SC motif sequence consists of a core of a branched, aliphatic residues in the first position and a critical lysine then glutamate in the second and fourth positions. Flanking this core is a proline or glycine residue located a few residues upstream and downstream of the SC core (Iñiguez-Lluhí and Pearce, 2000; Mukherjee et al., 2012). It was ultimately discovered that SC motifs align with the SUMOylation consensus motif and the lysine in the SC motif is indeed a SUMO target (Holmstrom et al., 2003; Poukka et al., 2000; Subramanian et al., 2003). Just as loss of the critical lysine in SC motifs abrogates their transcriptional repression *in vitro*, there are diseases related to the functional consequences of SC motif disruption.

The androgen receptor (AR) has two SC motifs in its A/B NTD that serve as SUMOylation sites (Figure 1-5). In Kennedy disease, also known as spinal and bulbar muscular atrophy (SBMA), there is a polyglutamine expansion within the AR NTD. This leads to misfolding, aggregation, and oligomerization of AR, causing lower motor neuronal toxicity. Our group has shown that SUMOylation of these polyQ ARs can limit the pathogenic AR aggregation, giving SUMO a role in disease beyond transcriptional effects (Mukherjee et al., 2009).

The proline and glycine amino acids that are upstream and downstream of both SC motifs in AR have been identified for mutation in numerous conditions implicated in AR biology, including degrees of androgen insensitivity syndrome (AIS) and associated oligospermia, infertility, and micropenis (Bhangoo et al., 2010; Ferlin et al., 2006; Hiort et al., 2000; Mukherjee et al., 2012). These mutations lead to reduced SUMOylation of AR and induction of endogenous AR target genes (Mukherjee et al., 2012). Mutations and

deletions in this flanking SC motif region and the SUMO targeted lysine have also been implicated in prostate cancer (Hyytinen et al., 2002; Mukherjee et al., 2012; Robinson et al., 2015).

7. SUMO and Prostate Cancer

Prostate cancer is the most common cancer diagnosed in men, making up nearly 20% of all male cancer diagnoses (Siegel et al., 2017). Although the rate of incidence is declining, in part due to more accurate diagnoses, approximately 1 out of every 8 men will be diagnosed with prostate cancer within his lifetime, creating a significant societal burden and resulting in it being the third leading cause of male cancer death.

At the center of prostate cancer is the androgen receptor. This nuclear receptor resides in the cytosol when inactive and binds androgens such as testosterone and its higher affinity ligand dihydrotestosterone (DHT) (Figure 1-6). Upon activation, it enters the nucleus via the nuclear localization signal found in the hinge region between its DBD and LBD. AR dimerizes and binds to androgen response elements (AREs) on the DNA to help activate transcription of genes essential to normal male sexual development and bone, muscle, and fat metabolism. In prostate cancer, this system is hijacked by oncogenes and AR is at the center of the disease. In 1941, Huggins and Hodges first discovered the role of androgen signaling in prostate cancer by demonstrating that surgical castration or androgen ablation therapy provided palliative effects in patients with metastatic prostate cancer (Huggins and Hodges, 1941; Huggins et al., 1941). This forms the basis of most prostate cancer therapeutics today, including androgen deprivation therapy (ADT), antiandrogen treatment, and inhibition of cytochrome p450 enzymes involved in steroid

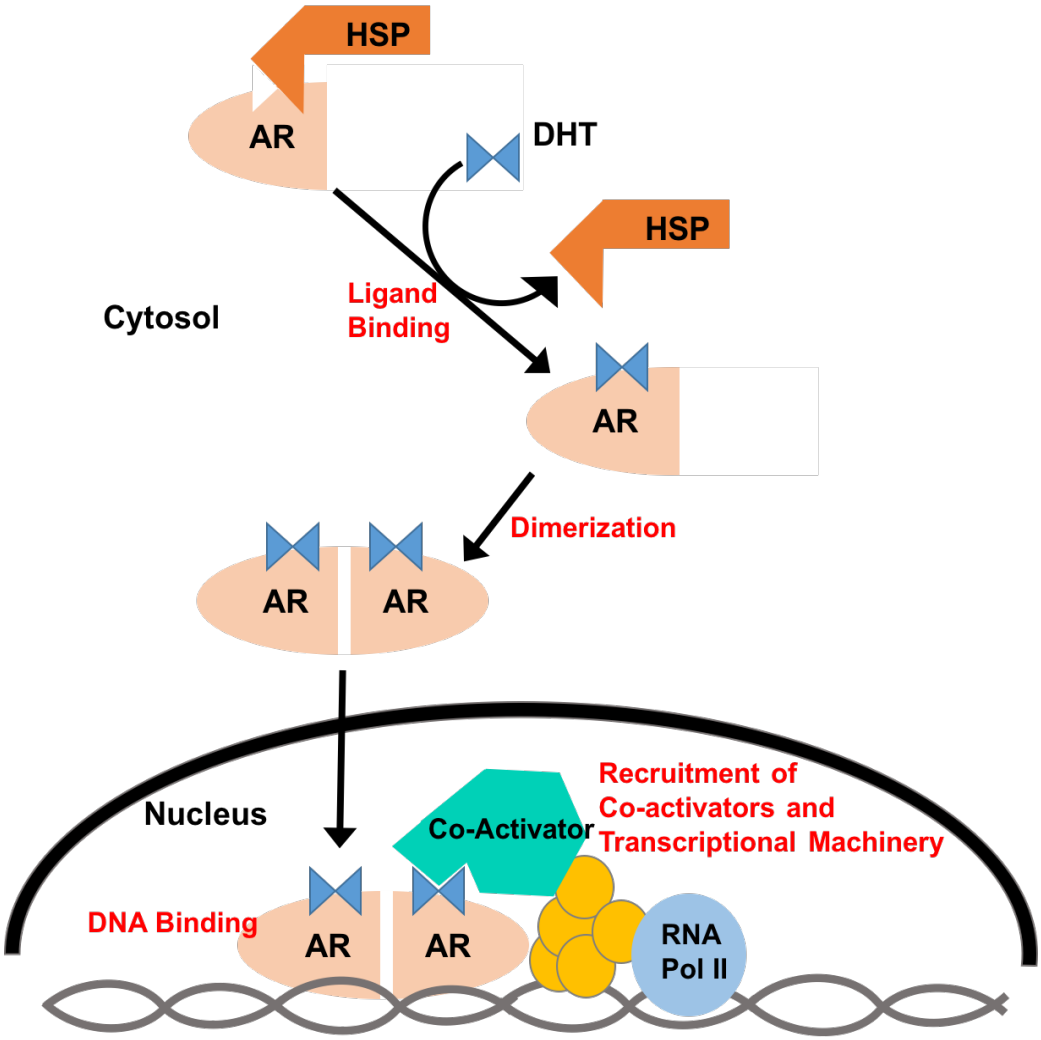


Figure 1-6 Androgen receptor signaling pathway.
The androgen receptor (AR) resides in the cytosol in complex with heat shock proteins (HSP) chaperones. Ligand binding by androgens, dihydrotestosterone (DHT) in this example, displaces the HSPs, leading to activation of AR. AR dimerizes and translocates to the nucleus where it binds to specific DNA sequences, termed androgen response elements (AREs). DNA binding leads to recruitment of AR co-regulators and proteins of the general transcription machinery, including RNA polymerase II (RNA Pol II), to initiate transcription of the respective DNA.

hormone synthesis. In ADT, gonadotropin releasing hormone (GnRH) agonists or antagonists are administered to eliminate GnRH pulses and instead generate a constant release or absence of the hormone (Tolis et al., 1982). This ultimately inhibits synthesis

of androgens in the testes, yielding castrate levels of circulating androgens by chemical means. Common AR antagonists occlude access to the LBD by androgen agonists to limit AR activation. Prostate cancer cells have also been shown to upregulate the enzymes involved in steroid hormone synthesis to allow for enhanced localized levels of androgens for increased AR signaling despite castrate levels in the serum (Cai and Balk, 2011; Montgomery et al., 2008; Nishiyama et al., 2004; Stanbrough et al., 2006; Titus et al., 2005). Abiraterone acetate is a third type of therapeutic for PC that targets cytochrome p450 17A1 to inhibit localized prostatic synthesis directly.

While these therapies listed above do provide initial involution of both healthy and cancerous prostatic tissue, approximately 80-90% of the patients relapse after 18-36 months through a variety of mechanisms in which AR has garnered the ability to maintain signaling despite the attempts to limit its activation (Chuu et al., 2011; Pienta and Bradley, 2006). These mechanisms include:

1. overexpression of AR to increase the receptor reserve and androgen sensitivity,
2. mutations in the LBD that create promiscuity in agonists to activate the receptor (both therapeutic antagonists and other steroid hormones),
3. alternative signal activation by other signaling proteins,
4. AR bypass by other TFs including glucocorticoid receptor activation which has a highly overlapping cistrome with AR, and
5. increased constitutive activity which can be a result of alternative splice forms of AR where the LBD is absent (Arora et al., 2013; Dehm et al., 2008; Guo et al., 2009; Hara et al., 2003; Hu et al., 2009; Sun et al., 2010; Suzuki et al., 1993; Tan

et al., 1997; Veldscholte et al., 1990; Visakorpi et al., 1995; Watson et al., 2015; Wijngaart et al., 2010; Zhao et al., 2000).

Despite all these resistance mechanisms, AR remains at the center of the proliferative drive of the disease in the majority of patients (Chandrasekar et al., 2015; Kahn et al., 2014). At this advanced stage, when the disease no longer appropriately responds to the three types of therapies identified above, it is identified as castration resistant prostate cancer (CRPC). While there are some therapies that exist beyond this stage, they are not curative and only temporarily extend survival by a few months. There is a need for new alternative therapies to address this deadly, advanced disease and targeting the SUMO pathway is among those efforts being employed.

As mentioned previously, SUMOylation of AR limits its transcriptional activity at compound AREs (Figure 1-5). With respect to the resistance mechanisms in CRPC listed above, AR SUMOylation is able to limit AR transcription in the context of both agonist activation and constitutive receptor activation, making the SUMO pathway an attractive target for novel therapeutics. Notably, unlike other nuclear receptors, the hormone-independent AF-1 of AR is stronger than its ligand induced AF-2 function in stimulating transcription (Ikonen et al., 1997; Moilanen et al., 1997). Existing clinical and biochemical data on prostate cancer further support that aberrations in the SUMO pathway that reduce AR SUMOylation are a way to leverage AR activity for cancer cell proliferation.

The AR is primarily modified by SUMO1 among the SUMO isoforms and SENP1 is the primary SUMO protease responsible for deSUMOylating AR (Kaikkonen et al., 2008). SENP1 is overexpressed in prostate cancer and can actually induce prostatic intraepithelial neoplasia, also known as PIN, a stage that precedes the development of

normal prostate epithelia to cancerous tissue (Bawa-Khalfe and Yeh, 2010; Bawa-Khalfe et al., 2010; Cheng et al., 2006). In addition, SENP1 expression is induced by AR activity as an androgen responsive gene (Bawa-Khalfe et al., 2007). With SENP1 as a positive modulator of AR activity via its deSUMOylation function, a positive feedback loop is generated between these two species to further the proliferative drive of the cancerous tissue, fostering prostate cancer progression and metastasis (Figure 1-7) (Wang et al., 2013). Interestingly, genes related to the cell cycle and cellular movement, death, proliferation, and development are differentially regulated between SUMOylated and non-SUMOylated AR, with the non-SUMOylated form showing an enrichment in these gene

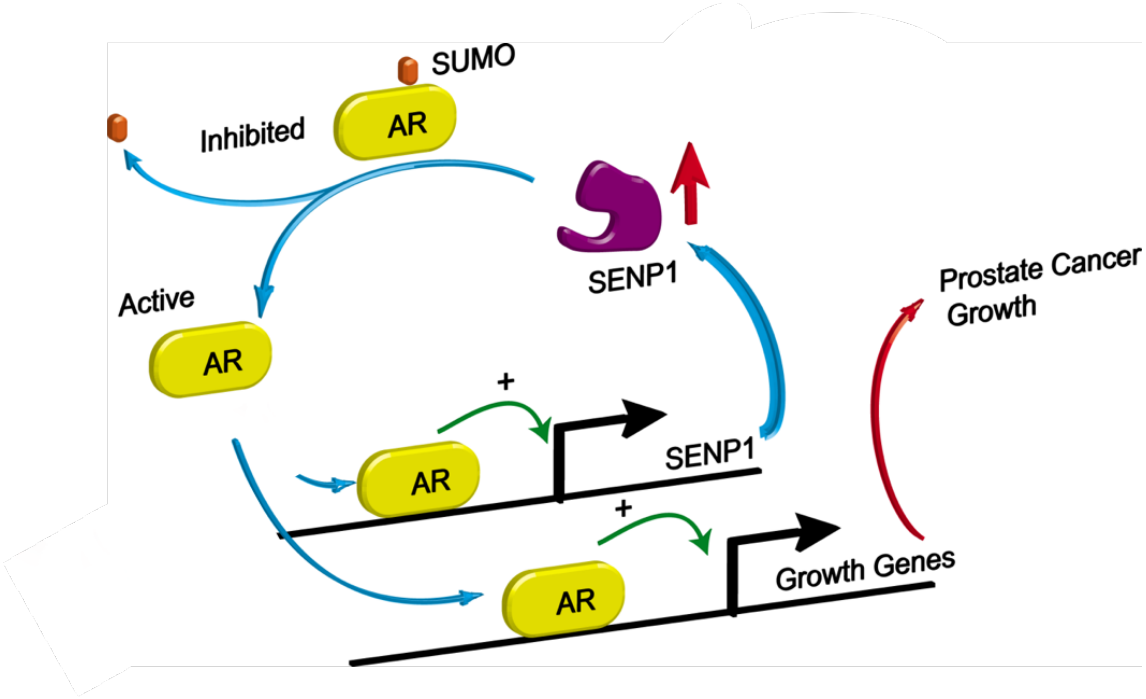


Figure 1-7 Summary of SUMO-AR-SENP1 interactions in prostate cancer.
When AR is SUMOylated (top) its transcriptional activity is inhibited, which limits AR's ability to activate transcription of genes involved in the growth of prostate cancer cells. Loss of SUMOylation on AR is performed by SENP1 and this generates the more active deSUMOylated form of AR to promote prostate cancer growth. One of the gene targets of AR is the gene for SENP1, creating a positive feedback loop where overexpression of SENP1 leads to enhanced AR activity, sustaining further SENP1 expression. Figure design by Jorge Iñiguez-Lluhí.

targets (Sutinen et al., 2014). In agreement with the role of overexpressed SENP1 in prostate cancer, knockdown of SENP1 via siRNA has been shown to attenuate expression of AR target genes while slowing prostate cancer cell growth and inducing senescence in cell culture (Kaikkonen et al., 2008). In patient cohorts, overexpression of SENP1 was associated with an array of characteristics indicating unfavorable disease outlook, including decreased biochemical recurrence free survival, increased tumor cell proliferation, higher Gleason score, positive ERG fusion status, PTEN deletion, and positive lymph node status (Burdelski et al., 2015; Li et al., 2013).

With strong clinical and biochemical target validation, SENP1 is an attractive drug target for new therapeutics in prostate cancer. In addition, crystal structures of the catalytic domains of the SENP family of enzymes allow for improved molecular modeling in drug development. Notably, with the SENP enzymes functioning as cysteine proteases, cysteine protease inhibition has been successfully approved for cancer therapy in terms of the proteasomal inhibitor bortezomib. In contrast to the current armamentarium of therapeutics, the SUMO1/SENP1 system functions at the NTD of AR, rather than through the LBD where existing observed mechanisms of resistance reside. Together, these attributes make SENP1 a novel, promising target for new prostate cancer therapeutics. Through high throughput screening and detailed kinetic characterization, the objective of this thesis work was to identify and study small molecule inhibitors specific to SENP1 for development of potential therapeutic scaffolds and probes to study the SUMO/SENP system and its regulation of activity.

8. References

1. Arora, V.K., Schenkein, E., Murali, R., Subudhi, S.K., Wongvipat, J., Balbas, M.D., Shah, N., Cai, L., Efstathiou, E., Logothetis, C., et al. (2013). Glucocorticoid Receptor Confers Resistance to Antiandrogens by Bypassing Androgen Receptor Blockade. Cell *155*, 1309–1322.

2. Bailey, D., and O'Hare, P. (2004). Characterization of the Localization and Proteolytic Activity of the SUMO-specific Protease, SENP1. J. Biol. Chem. *279*, 692–703.

3. Bawa-Khalfe, T., and Yeh, E.T.H. (2010). SUMO Losing Balance. Genes Cancer *1*, 748–752.

4. Bawa-Khalfe, T., Cheng, J., Wang, Z., and Yeh, E.T.H. (2007). Induction of the SUMO-specific Protease 1 Transcription by the Androgen Receptor in Prostate Cancer Cells. J. Biol. Chem. *282*, 37341–37349.

5. Bawa-Khalfe, T., Cheng, J., Lin, S.-H., Ittmann, M.M., and Yeh, E.T.H. (2010). SENP1 Induces Prostatic Intraepithelial Neoplasia through Multiple Mechanisms. J. Biol. Chem. *285*, 25859–25866.

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

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

8. Berta, M.A., Mazure, N., Hattab, M., Pouysségur, J., and Brahimi-Horn, M.C. (2007). SUMOylation of hypoxia-inducible factor-1α reduces its transcriptional activity. Biochem. Biophys. Res. Commun. *360*, 646–652.

9. Bhangoo, A., Paris, F., Philibert, P., Audran, F., Ten, S., and Sultan, C. (2010). Isolated micropenis reveals partial androgen insensitivity syndrome confirmed by molecular analysis. Asian J. Androl. *12*, 561–566.

10. Boddy, M.N., Howe, K., Etkin, L.D., Solomon, E., and Freemont, P.S. (1996). PIC 1, a novel ubiquitin-like protein which interacts with the PML component of a multiprotein complex that is disrupted in acute promyelocytic leukaemia. Oncogene *13*, 971–982.

11. Bohren, K.M., Nadkarni, V., Song, J.H., Gabbay, K.H., and Owerbach, D. (2004). A M55V Polymorphism in a Novel SUMO Gene (SUMO-4) Differentially Activates Heat Shock Transcription Factors and Is Associated with Susceptibility to Type I Diabetes Mellitus. J. Biol. Chem. *279*, 27233–27238.

12. Burdelski, C., Menan, D., Tsourlakis, M.C., Kluth, M., Hube-Magg, C., Melling, N., Minner, S., Koop, C., Graefen, M., Heinzer, H., et al. (2015). The prognostic value of SUMO1/Sentrin specific peptidase 1 (SENP1) in prostate cancer is limited to ERG-fusion positive tumors lacking PTEN deletion. BMC Cancer *15*, 538.

13. Cai, C., and Balk, S.P. (2011). Intratumoral androgen biosynthesis in prostate cancer pathogenesis and response to therapy. Endocr. Relat. Cancer *18*, R175–R182.

14. Cappadocia, L., and Lima, C.D. (2017). Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. Chem. Rev.

15. Chandrasekar, T., Yang, J.C., Gao, A.C., and Evans, C.P. (2015). Mechanisms of resistance in castration-resistant prostate cancer (CRPC). Transl. Androl. Urol. *4*, 365–380.

16. Chang, C.-C., Naik, M.T., Huang, Y.-S., Jeng, J.-C., Liao, P.-H., Kuo, H.-Y., Ho, C.-C., Hsieh, Y.-L., Lin, C.-H., Huang, N.-J., et al. (2011). Structural and Functional Roles of Daxx SIM Phosphorylation in SUMO Paralog-Selective Binding and Apoptosis Modulation. Mol. Cell *42*, 62–74.

17. Chen, A., Mannen, H., and Li, S.S. (1998). Characterization of mouse ubiquitin-like SMT3A and SMT3B cDNAs and gene/pseudogenes. Biochem. Mol. Biol. Int. *46*, 1161–1174.

18. Chen, T., Zhou, T., He, B., Yu, H., Guo, X., Song, X., and Sha, J. (2014). mUbiSiDa: A Comprehensive Database for Protein Ubiquitination Sites in Mammals. PLOS ONE *9*, e85744.

19. Cheng, J., Bawa, T., Lee, P., Gong, L., and Yeh, E.T.H. (2006). Role of Desumoylation in the Development of Prostate Cancer. Neoplasia N. Y. N *8*, 667–676.

20. Cheng, J., Kang, X., Zhang, S., and Yeh, E.T.H. (2007). SUMO-Specific Protease 1 Is Essential for Stabilization of HIF1α during Hypoxia. Cell *131*, 584–595.

21. Chuu, C.-P., Kokontis, J.M., Hiipakka, R.A., Fukuchi, J., Lin, H.-P., Lin, C.-Y., Huo, C., and Su, L.-C. (2011). Androgens as therapy for androgen receptor-positive castration-resistant prostate cancer. J. Biomed. Sci. *18*, 63.

22. Chymkowitch, P., Nguéa P, A., and Enserink, J.M. (2015). SUMO-regulated transcription: Challenging the dogma. BioEssays *37*, 1095–1105.

23. Cox, B., Briscoe, J., and Ulloa, F. (2010). SUMOylation by Pias1 Regulates the Activity of the Hedgehog Dependent Gli Transcription Factors. PLOS ONE *5*, e11996.

24. Cremona, C.A., Sarangi, P., Yang, Y., Hang, L.E., Rahman, S., and Zhao, X. (2012). Extensive DNA Damage-Induced Sumoylation Contributes to Replication and Repair and Acts in Addition to the Mec1 Checkpoint. Mol. Cell *45*, 422–432.

25. Danciu, T.E., Chupreta, S., Cruz, O., Fox, J.E., Whitman, M., and Iñiguez-Lluhí, J.A. (2012). Small Ubiquitin-like Modifier (SUMO) Modification Mediates Function of the Inhibitory Domains of Developmental Regulators FOXC1 and FOXC2. J. Biol. Chem. *287*, 18318–18329.

26. Dehm, S.M., Schmidt, L.J., Heemers, H.V., Vessella, R.L., and Tindall, D.J. (2008). Splicing of a Novel Androgen Receptor Exon Generates a Constitutively Active Androgen Receptor that Mediates Prostate Cancer Therapy Resistance. Cancer Res. *68*, 5469–5477.

27. Desterro, J.M.P., Rodriguez, M.S., and Hay, R.T. (1998). SUMO-1 Modification of IκBα Inhibits NF-κB Activation. Mol. Cell *2*, 233–239.

28. Desterro, J.M.P., Rodriguez, M.S., Kemp, G.D., and Hay, R.T. (1999). Identification of the Enzyme Required for Activation of the Small Ubiquitin-like Protein SUMO-1. J. Biol. Chem. *274*, 10618–10624.

29. Dou, H., Huang, C., Singh, M., Carpenter, P.B., and Yeh, E.T.H. (2010). Regulation of DNA Repair through DeSUMOylation and SUMOylation of Replication Protein A Complex. Mol. Cell *39*, 333–345.

30. Eisenhardt, N., Chaugule, V.K., Koidl, S., Droescher, M., Dogan, E., Rettich, J., Sutinen, P., Imanishi, S.Y., Hofmann, K., Palvimo, J.J., et al. (2015). A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. Nat. Struct. Mol. Biol. *22*, 959–967.

31. Elsasser, S., and Finley, D. (2005). Delivery of ubiquitinated substrates to protein-unfolding machines. Nat. Cell Biol. *7*, 742–749.

32. Evdokimov, E., Sharma, P., Lockett, S.J., Lualdi, M., and Kuehn, M.R. (2008). Loss of SUMO1 in mice affects RanGAP1 localization and formation of PML nuclear bodies, but is not lethal as it can be compensated by SUMO2 or SUMO3. J Cell Sci *121*, 4106–4113.

33. Ferlin, A., Vinanzi, C., Garolla, A., Selice, R., Zuccarello, D., Cazzadore, C., and Foresta, C. (2006). Male infertility and androgen receptor gene mutations: clinical features and identification of seven novel mutations. Clin. Endocrinol. (Oxf.) *65*, 606–610.

34. Galanty, Y., Belotserkovskaya, R., Coates, J., and Jackson, S.P. (2012). RNF4, a SUMO-targeted ubiquitin E3 ligase, promotes DNA double-strand break repair. Genes Dev. *26*, 1179–1195.

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

36. Germain, P., Staels, B., Dacquet, C., Spedding, M., and Laudet, V. (2006). Overview of Nomenclature of Nuclear Receptors. Pharmacol. Rev. *58*, 685–704.

37. Gill, G. (2005). Something about SUMO inhibits transcription. Curr. Opin. Genet. Dev. *15*, 536–541.

38. Girdwood, D., Bumpass, D., Vaughan, O.A., Thain, A., Anderson, L.A., Snowden, A.W., Garcia-Wilson, E., Perkins, N.D., and Hay, R.T. (2003). p300 Transcriptional Repression Is Mediated by SUMO Modification. Mol. Cell *11*, 1043–1054.

39. Gong, L., Li, B., Millas, S., and Yeh, E.T.H. (1999). Molecular cloning and characterization of human AOS1 and UBA2, components of the sentrin-activating enzyme complex. FEBS Lett. *448*, 185–189.

40. Gong, L., Millas, S., Maul, G.G., and Yeh, E.T.H. (2000). Differential Regulation of Sentrinized Proteins by a Novel Sentrin-specific Protease. J. Biol. Chem. *275*, 3355–3359.

41. Guo, D., Li, M., Zhang, Y., Yang, P., Eckenrode, S., Hopkins, D., Zheng, W., Purohit, S., Podolsky, R.H., Muir, A., et al. (2004). A functional variant of SUMO4, a new IκBα modifier, is associated with type 1 diabetes. Nat. Genet. *36*, 837–841.

42. Guo, Z., Yang, X., Sun, F., Jiang, R., Linn, D.E., Chen, H., Chen, H., Kong, X., Melamed, J., Tepper, C.G., et al. (2009). A Novel Androgen Receptor Splice Variant Is Up-regulated during Prostate Cancer Progression and Promotes Androgen Depletion–Resistant Growth. Cancer Res. *69*, 2305–2313.

43. Guzzo, C.M., Berndsen, C.E., Zhu, J., Gupta, V., Datta, A., Greenberg, R.A., Wolberger, C., and Matunis, M.J. (2012). RNF4-Dependent Hybrid SUMO-Ubiquitin Chains Are Signals for RAP80 and Thereby Mediate the Recruitment of BRCA1 to Sites of DNA Damage. Sci Signal *5*, ra88-ra88.

44. Hang, J., and Dasso, M. (2002). Association of the Human SUMO-1 Protease SENP2 with the Nuclear Pore. J. Biol. Chem. *277*, 19961–19966.

45. Hannich, J.T., Lewis, A., Kroetz, M.B., Li, S.-J., Heide, H., Emili, A., and Hochstrasser, M. (2005). Defining the SUMO-modified Proteome by Multiple Approaches in Saccharomyces cerevisiae. J. Biol. Chem. *280*, 4102–4110.

46. Hanpude, P., Bhattacharya, S., Dey, A.K., and Maiti, T.K. (2015). Deubiquitinating enzymes in cellular signaling and disease regulation. IUBMB Life *67*, 544–555.

47. Hara, T., Miyazaki, J., Araki, H., Yamaoka, M., Kanzaki, N., Kusaka, M., and Miyamoto, M. (2003). Novel Mutations of Androgen Receptor: A Possible Mechanism of Bicalutamide Withdrawal Syndrome. Cancer Res. *63*, 149–153.

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

49. Hecker, C.-M., Rabiller, M., Haglund, K., Bayer, P., and Dikic, I. (2006). Specification of SUMO1- and SUMO2-interacting Motifs. J. Biol. Chem. *281*, 16117–16127.

50. Hendriks, I.A., and Vertegaal, A.C.O. (2016). A comprehensive compilation of SUMO proteomics. Nat. Rev. Mol. Cell Biol. *17*, 581–595.

51. Hendriks, I.A., Schimmel, J., Eifler, K., Olsen, J.V., and Vertegaal, A.C.O. (2015). Ubiquitin-specific Protease 11 (USP11) Deubiquitinates Hybrid Small Ubiquitin-like Modifier (SUMO)-Ubiquitin Chains to Counteract RING Finger Protein 4 (RNF4). J. Biol. Chem. *290*, 15526–15537.

52. Hendriks, I.A., Lyon, D., Young, C., Jensen, L.J., Vertegaal, A.C.O., and Nielsen, M.L. (2017). Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. Nat. Struct. Mol. Biol. *24*, 325–336.

53. Hershko, A., Heller, H., Elias, S., and Ciechanover, A. (1983). Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. J. Biol. Chem. *258*, 8206–8214.

54. Hickey, C.M., Wilson, N.R., and Hochstrasser, M. (2012). Function and regulation of SUMO proteases. Nat. Rev. Mol. Cell Biol. *13*, 755–766.

55. Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A., and Sistonen, L. (2006). PDSM, a motif for phosphorylation-dependent SUMO modification. Proc. Natl. Acad. Sci. U. S. A. *103*, 45–50.

56. Hiort, O., Holterhus, P.-M., Horter, T., Schulze, W., Kremke, B., Bals-Pratsch, M., Sinnecker, G.H.G., and Kruse, K. (2000). Significance of Mutations in the Androgen Receptor Gene in Males with Idiopathic Infertility. J. Clin. Endocrinol. Metab. *85*, 2810–2815.

57. Holmstrom, S., Antwerp, M.E.V., and Iñiguez-Lluhí, J.A. (2003). Direct and distinguishable inhibitory roles for SUMO isoforms in the control of transcriptional synergy. Proc. Natl. Acad. Sci. *100*, 15758–15763.

58. Hu, R., Dunn, T.A., Wei, S., Isharwal, S., Veltri, R.W., Humphreys, E., Han, M., Partin, A.W., Vessella, R.L., Isaacs, W.B., et al. (2009). Ligand-Independent Androgen Receptor Variants Derived from Splicing of Cryptic Exons Signify Hormone-Refractory Prostate Cancer. Cancer Res. *69*, 16–22.

59. Huang, T.T., and D'Andrea, A.D. (2006). Regulation of DNA repair by ubiquitylation. Nat. Rev. Mol. Cell Biol. *7*, 323–334.

60. Huang, W.-C., Ko, T.-P., Li, S.S.-L., and Wang, A.H.-J. (2004). Crystal structures of the human SUMO-2 protein at 1.6 Å and 1.2 Å resolution. Eur. J. Biochem. *271*, 4114–4122.

61. Huggins, C., and Hodges, C.V. (1941). Studies on prostatic cancer: I. The effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate. CA. Cancer J. Clin. *22*, 232–240.

62. Huggins, C., Stevens, R.E., and Hodges, C.V. (1941). STUDIES ON PROSTATIC CANCER: II. THE EFFECTS OF CASTRATION ON ADVANCED CARCINOMA OF THE PROSTATE GLAND. Arch. Surg. *43*, 209–223.

63. Hurley, J.H., Lee, S., and Prag, G. (2006). Ubiquitin-binding domains. Biochem. J. *399*, 361–372.

64. Hyytinen, E.-R., Haapala, K., Thompson, J., Lappalainen, I., Roiha, M., Rantala, I., Helin, H.J., Jänne, O.A., Vihinen, M., Palvimo, J.J., et al. (2002). Pattern of Somatic Androgen Receptor Gene Mutations in Patients with Hormone-Refractory Prostate Cancer. Lab. Invest. *82*, 1591–1598.

65. Ikonen, T., Palvimo, J.J., and Jänne, O.A. (1997). Interaction between the Amino- and Carboxyl-terminal Regions of the Rat Androgen Receptor Modulates Transcriptional Activity and Is Influenced by Nuclear Receptor Coactivators. J. Biol. Chem. *272*, 29821–29828.

66. Iñiguez-Lluhí, J.A., and Pearce, D. (2000). A Common Motif within the Negative Regulatory Regions of Multiple Factors Inhibits Their Transcriptional Synergy. Mol. Cell. Biol. *20*, 6040–6050.

67. Itahana, Y., Yeh, E.T.H., and Zhang, Y. (2006). Nucleocytoplasmic Shuttling Modulates Activity and Ubiquitination-Dependent Turnover of SUMO-Specific Protease 2. Mol. Cell. Biol. *26*, 4675–4689.

68. Jiang, J. (2006). Regulation of Hh/Gli Signaling by Dual Ubiquitin Pathways. Cell Cycle *5*, 2457–2463.

69. Jin, J., Li, X., Gygi, S.P., and Harper, J.W. (2007). Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. Nature *447*, 1135–1138.

70. Johnson, E.S., and Blobel, G. (1997). Ubc9p Is the Conjugating Enzyme for the Ubiquitin-like Protein Smt3p. J. Biol. Chem. *272*, 26799–26802.

71. Kahn, B., Collazo, J., and Kyprianou, N. (2014). Androgen Receptor as a Driver of Therapeutic Resistance in Advanced Prostate Cancer. Int. J. Biol. Sci. *10*, 588–595.

72. Kahyo, T., Nishida, T., and Yasuda, H. (2001). Involvement of PIAS1 in the Sumoylation of Tumor Suppressor p53. Mol. Cell *8*, 713–718.

73. Kaikkonen, S., Jaaskelainen, T., Karvonen, U., Rytinki, M.M., Makkonen, H., Gioeli, D., Paschal, B.M., and Palvimo, J.J. (2008). SUMO-Specific Protease 1 (SENP1) Reverses the Hormone-Augmented SUMOylation of Androgen Receptor and Modulates Gene Responses in Prostate Cancer Cells. Mol. Endocrinol. *23*, 292–307.

74. Kamitani, T., Kito, K., Nguyen, H.P., Fukuda-Kamitani, T., and Yeh, E.T.H. (1998). Characterization of a Second Member of the Sentrin Family of Ubiquitin-like Proteins. J. Biol. Chem. *273*, 11349–11353.

75. Kang, X., Qi, Y., Zuo, Y., Wang, Q., Zou, Y., Schwartz, R.J., Cheng, J., and Yeh, E.T.H. (2010). SUMO-Specific Protease 2 Is Essential for Suppression of Polycomb Group Protein-Mediated Gene Silencing during Embryonic Development. Mol. Cell *38*, 191–201.

76. Kaur, K., Park, H., Pandey, N., Azuma, Y., and Guzman, R.N.D. (2017). Identification of a new small ubiquitin-like modifier (SUMO)-interacting motif in the E3 ligase PIASy. J. Biol. Chem. *292*, 10230–10238.

77. Kerscher, O. (2007). SUMO junction--what's your function? New insights through SUMO-interacting motifs. EMBO Rep. *8*, 550–555.

78. Keusekotten, K., Bade, V.N., Meyer-Teschendorf, K., Sriramachandran, A.M., Fischer-Schrader, K., Krause, A., Horst, C., Schwarz, G., Hofmann, K., Dohmen, R.J., et al. (2014). Multivalent interactions of the SUMO-interaction motifs in RING finger protein 4 determine the specificity for chains of the SUMO. Biochem. J. *457*, 207–214.

79. Khoury, G.A., Baliban, R.C., and Floudas, C.A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. Sci. Rep. *1*.

80. Kim, K.I., Baek, S.H., Jeon, Y.-J., Nishimori, S., Suzuki, T., Uchida, S., Shimbara, N., Saitoh, H., Tanaka, K., and Chung, C.H. (2000). A New SUMO-1-specific Protease, SUSP1, That Is Highly Expressed in Reproductive Organs. J. Biol. Chem. *275*, 14102–14106.

81. Kolli, N., Mikolajczyk, J., Drag, M., Mukhopadhyay, D., Moffatt, N., Dasso, M., Salvesen, G., and Wilkinson, K.D. (2010). Distribution and paralogue specificity of mammalian deSUMOylating enzymes. Biochem. J. *430*, 335–344.

82. Komander, D., and Rape, M. (2012). The Ubiquitin Code. Annu. Rev. Biochem. *81*, 203–229.

83. Komander, D., Clague, M.J., and Urbé, S. (2009). Breaking the chains: structure and function of the deubiquitinases. Nat. Rev. Mol. Cell Biol. *10*, 550–563.

84. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG. Bioinformatics *2007 23(21):* 2947-2948. doi:10.1093/bioinformatics/btm404.

85. Lecona, E., Rodriguez-Acebes, S., Specks, J., Lopez-Contreras, A.J., Ruppen, I., Murga, M., Muñoz, J., Mendez, J., and Fernandez-Capetillo, O. (2016). USP7 is a SUMO deubiquitinase essential for DNA replication. Nat. Struct. Mol. Biol. *23*, 270–277.

86. Li, T., Huang, S., Dong, M., Gui, Y., and Wu, D. (2013). Prognostic impact of SUMO-specific protease 1 (SENP1) in prostate cancer patients undergoing radical prostatectomy. Urol. Oncol. Semin. Orig. Investig. *31*, 1539–1545.

87. Lyst, M.J., and Stancheva, I. (2007). A role for SUMO modification in transcriptional repression and activation. Biochem. Soc. Trans. *35*, 1389–1392.

88. Matic, I., Schimmel, J., Hendriks, I.A., van Santen, M.A., van de Rijke, F., van Dam, H., Gnad, F., Mann, M., and Vertegaal, A.C.O. (2010). Site-Specific Identification of SUMO-2 Targets in Cells Reveals an Inverted SUMOylation Motif and a Hydrophobic Cluster SUMOylation Motif. Mol. Cell *39*, 641–652.

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

90. Mendes, A.V., Grou, C.P., Azevedo, J.E., and Pinto, M.P. (2016). Evaluation of the activity and substrate specificity of the human SENP family of SUMO proteases. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1863*, 139–147.

91. Mikolajczyk, J., Drag, M., Békés, M., Cao, J.T., Ronai, Z. 'ev, and Salvesen, G.S. (2007). Small Ubiquitin-related Modifier (SUMO)-specific Proteases PROFILING THE SPECIFICITIES AND ACTIVITIES OF HUMAN SENPs. J. Biol. Chem. *282*, 26217–26224.

92. Minty, A., Dumont, X., Kaghad, M., and Caput, D. (2000). Covalent Modification of p73α by SUMO-1 TWO-HYBRID SCREENING WITH p73 IDENTIFIES NOVEL SUMO-1-INTERACTING PROTEINS AND A SUMO-1 INTERACTION MOTIF. J. Biol. Chem. *275*, 36316–36323.

93. Moilanen, A., Rouleau, N., Ikonen, T., Palvimo, J.J., and Jänne, O.A. (1997). The presence of a transcription activation function in the hormone-binding domain of androgen receptor is revealed by studies in yeast cells. FEBS Lett. *412*, 355–358.

94. Montgomery, R.B., Mostaghel, E.A., Vessella, R., Hess, D.L., Kalhorn, T.F., Higano, C.S., True, L.D., and Nelson, P.S. (2008). Maintenance of Intratumoral Androgens in Metastatic Prostate Cancer: A Mechanism for Castration-Resistant Tumor Growth. Cancer Res. *68*, 4447–4454.

95. Morreale, F.E., and Walden, H. (2016). Types of Ubiquitin Ligases. Cell *165*, 248–248.e1.

96. Mueller, T., Breuer, P., Schmitt, I., Walter, J., Evert, B.O., and Wüllner, U. (2009). CK2-dependent phosphorylation determines cellular localization and stability of ataxin-3. Hum. Mol. Genet. *18*, 3334–3343.

97. Mukherjee, S., Thomas, M., Dadgar, N., Lieberman, A.P., and Iñiguez-Lluhí, J.A. (2009). Small Ubiquitin-like Modifier (SUMO) Modification of the Androgen Receptor Attenuates Polyglutamine-mediated Aggregation. J. Biol. Chem. *284*, 21296–21306.

98. Mukherjee, S., Cruz-Rodríguez, O., Bolton, E., and Iñiguez-Lluhí, J.A. (2012). The in Vivo Role of Androgen Receptor SUMOylation as Revealed by Androgen Insensitivity Syndrome and Prostate Cancer Mutations Targeting the Proline/Glycine Residues of Synergy Control Motifs. J. Biol. Chem. *287*, 31195–31206.

99. Mukhopadhyay, D., and Dasso, M. (2007). Modification in reverse: the SUMO proteases. Trends Biochem. Sci. *32*, 286–295.

100. Nacerddine, K., Lehembre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P.P., and Dejean, A. (2005). The SUMO Pathway Is Essential for Nuclear Integrity and Chromosome Segregation in Mice. Dev. Cell *9*, 769–779.

101. Namanja, A.T., Li, Y.-J., Su, Y., Wong, S., Lu, J., Colson, L.T., Wu, C., Li, S.S.C., and Chen, Y. (2012). Insights into High Affinity Small Ubiquitin-like Modifier (SUMO) Recognition by SUMO-interacting Motifs (SIMs) Revealed by a Combination of NMR and Peptide Array Analysis. J. Biol. Chem. *287*, 3231–3240.

102. Nayak, A., and Müller, S. (2014). SUMO-specific proteases/isopeptidases: SENPs and beyond. Genome Biol. *15*, 422.

103. Nishida, T., Tanaka, H., and Yasuda, H. (2000). A novel mammalian Smt3-specific isopeptidase 1 (SMT3IP1) localized in the nucleolus at interphase. Eur. J. Biochem. *267*, 6423–6427.

104. Nishida, T., Kaneko, F., Kitagawa, M., and Yasuda, H. (2001). Characterization of a Novel Mammalian SUMO-1/Smt3-specific Isopeptidase, a Homologue of Rat Axam, Which Is an Axin-binding Protein Promoting β-Catenin Degradation. J. Biol. Chem. *276*, 39060–39066.

105. Nishiyama, T., Hashimoto, Y., and Takahashi, K. (2004). The Influence of Androgen Deprivation Therapy on Dihydrotestosterone Levels in the Prostatic Tissue of Patients with Prostate Cancer. Clin. Cancer Res. *10*, 7121–7126.

106. Ohuchi, T., Seki, M., Branzei, D., Maeda, D., Ui, A., Ogiwara, H., Tada, S., and Enomoto, T. (2008). Rad52 sumoylation and its involvement in the efficient induction of homologous recombination. DNA Repair *7*, 879–889.

107. Okuma, T., Honda, R., Ichikawa, G., Tsumagari, N., and Yasuda, H. (1999). In VitroSUMO-1 Modification Requires Two Enzymatic Steps, E1 and E2. Biochem. Biophys. Res. Commun. *254*, 693–698.

108. Okura, T., Gong, L., Kamitani, T., Wada, T., Okura, I., Wei, C.F., Chang, H.M., and Yeh, E.T. (1996). Protection against Fas/APO-1- and tumor necrosis factor-mediated cell death by a novel protein, sentrin. J. Immunol. *157*, 4277–4281.

109. Owerbach, D., McKay, E.M., Yeh, E.T.H., Gabbay, K.H., and Bohren, K.M. (2005). A proline-90 residue unique to SUMO-4 prevents maturation and sumoylation. Biochem. Biophys. Res. Commun. *337*, 517–520.

110. Peng, J., Schwartz, D., Elias, J.E., Thoreen, C.C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S.P. (2003). A proteomics approach to understanding protein ubiquitination. Nat. Biotechnol. *21*, 921–926.

111. Pichler, A., Gast, A., Seeler, J.S., Dejean, A., and Melchior, F. (2002). The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity. Cell *108*, 109–120.

112. Pienta, K.J., and Bradley, D. (2006). Mechanisms Underlying the Development of Androgen-Independent Prostate Cancer. Clin. Cancer Res. *12*, 1665–1671.

113. Poukka, H., Karvonen, U., Jänne, O.A., and Palvimo, J.J. (2000). Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). Proc. Natl. Acad. Sci. *97*, 14145–14150.

114. Pourcet, B., Pineda-Torra, I., Derudas, B., Staels, B., and Glineur, C. (2010). SUMOylation of Human Peroxisome Proliferator-activated Receptor α Inhibits Its Trans-activity through the Recruitment of the Nuclear Corepressor NCoR. J. Biol. Chem. *285*, 5983–5992.

115. Psakhye, I., and Jentsch, S. (2012). Protein Group Modification and Synergy in the SUMO Pathway as Exemplified in DNA Repair. Cell *151*, 807–820.

116. Robinson, D., Van Allen, E.M., Wu, Y.-M., Schultz, N., Lonigro, R.J., Mosquera, J.-M., Montgomery, B., Taplin, M.-E., Pritchard, C.C., Attard, G., et al. (2015). Integrative Clinical Genomics of Advanced Prostate Cancer. Cell *161*, 1215–1228.

117. Rodriguez, M.S., Dargemont, C., and Hay, R.T. (2001). SUMO-1 Conjugation in Vivo Requires Both a Consensus Modification Motif and Nuclear Targeting. J. Biol. Chem. *276*, 12654–12659.

118. Rosonina, E., Akhter, A., Dou, Y., Babu, J., and Theivakadadcham, V.S.S. (2017). Regulation of transcription factors by sumoylation. Transcription *0*, e1311829.

119. Rytinki, M.M., and Palvimo, J.J. (2008). SUMOylation Modulates the Transcription Repressor Function of RIP140. J. Biol. Chem. *283*, 11586–11595.

120. Rytinki, M.M., and Palvimo, J.J. (2009). SUMOylation attenuates the function of PGC-1alpha. J. Biol. Chem. jbc.M109.038943.

121. Saeki, Y., Sone, T., Toh-e, A., and Yokosawa, H. (2002). Identification of ubiquitin-like protein-binding subunits of the 26S proteasome. Biochem. Biophys. Res. Commun. *296*, 813–819.

122. Sampson, D.A., Wang, M., and Matunis, M.J. (2001). The Small Ubiquitin-like Modifier-1 (SUMO-1) Consensus Sequence Mediates Ubc9 Binding and Is Essential for SUMO-1 Modification. J. Biol. Chem. *276*, 21664–21669.

123. Schlesinger, D.H., Goldstein, G., and Niall, H.D. (1975). Complete amino acid sequence of ubiquitin, an adenylate cyclase stimulating polypeptide probably universal in living cells. Biochemistry (Mosc.) *14*, 2214–2218.

124. Schrader, E.K., Harstad, K.G., and Matouschek, A. (2009). Targeting proteins for degradation. Nat. Chem. Biol. *5*, 815–822.

125. Sharma, P., Yamada, S., Lualdi, M., Dasso, M., and Kuehn, M.R. (2013). Senp1 Is Essential for Desumoylating Sumo1-Modified Proteins but Dispensable for Sumo2 and Sumo3 Deconjugation in the Mouse Embryo. Cell Rep. *3*, 1640–1650.

126. Shen, Z., Pardington-Purtymun, P.E., Comeaux, J.C., Moyzis, R.K., and Chen, D.J. (1996). UBL1, a Human Ubiquitin-like Protein Associating with Human RAD51/RAD52 Proteins. Genomics *36*, 271–279.

127. Siegel, R.L., Miller, K.D., and Jemal, A. (2017). Cancer statistics, 2017. CA. Cancer J. Clin. *67*, 7–30.

128. Smet-Nocca, C., Wieruszeski, J.-M., Léger, H., Eilebrecht, S., and Benecke, A. (2011). SUMO-1 regulates the conformational dynamics of Thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. BMC Biochem. *12*, 4.

129. Sokratous, K., Hadjisavvas, A., Diamandis, E.P., and Kyriacou, K. (2014). The role of ubiquitin-binding domains in human pathophysiology. Crit. Rev. Clin. Lab. Sci. *51*, 280–290.

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

131. Song, J., Zhang, Z., Hu, W., and Chen, Y. (2005). Small Ubiquitin-like Modifier (SUMO) Recognition of a SUMO Binding Motif A REVERSAL OF THE BOUND ORIENTATION. J. Biol. Chem. *280*, 40122–40129.

132. Stanbrough, M., Bubley, G.J., Ross, K., Golub, T.R., Rubin, M.A., Penning, T.M., Febbo, P.G., and Balk, S.P. (2006). Increased Expression of Genes Converting

Adrenal Androgens to Testosterone in Androgen-Independent Prostate Cancer. Cancer Res. *66*, 2815–2825.

133. Steinacher, R., and Schär, P. (2005). Functionality of Human Thymine DNA Glycosylase Requires SUMO-Regulated Changes in Protein Conformation. Curr. Biol. *15*, 616–623.

134. Subramanian, L., Benson, M.D., and Iñiguez-Lluhí, J.A. (2003). A Synergy Control Motif within the Attenuator Domain of CCAAT/Enhancer-binding Protein α Inhibits Transcriptional Synergy through Its PIASy-enhanced Modification by SUMO-1 or SUMO-3. J. Biol. Chem. *278*, 9134–9141.

135. Sun, S., Sprenger, C.C.T., Vessella, R.L., Haugk, K., Soriano, K., Mostaghel, E.A., Page, S.T., Coleman, I.M., Nguyen, H.M., Sun, H., et al. (2010). Castration resistance in human prostate cancer is conferred by a frequently occurring androgen receptor splice variant. J. Clin. Invest. *120*, 2715–2730.

136. Sutinen, P., Malinen, M., Heikkinen, S., and Palvimo, J.J. (2014). SUMOylation modulates the transcriptional activity of androgen receptor in a target gene and pathway selective manner. Nucleic Acids Res. *42*, 8310–8319.

137. Suzuki, H., Sato, N., Watabe, Y., Masai, M., Seino, S., and Shimazaki, J. (1993). Androgen receptor gene mutations in human prostate cancer. J. Steroid Biochem. Mol. Biol. *46*, 759–765.

138. Tan, J., Sharief, Y., Hamil, K.G., Gregory, C.W., Zang, D.-Y., Sar, M., Gumerlock, P.H., White, deVere, W, R., Pretlow, T.G., et al. (1997). Dehydroepiandrosterone Activates Mutant Androgen Receptors Expressed in the Androgen-Dependent Human Prostate Cancer Xenograft CWR22 and LNCaP Cells. Mol. Endocrinol. *11*, 450–459.

139. Tatham, M.H., Jaffray, E., Vaughan, O.A., Desterro, J.M.P., Botting, C.H., Naismith, J.H., and Hay, R.T. (2001). Polymeric Chains of SUMO-2 and SUMO-3 Are Conjugated to Protein Substrates by SAE1/SAE2 and Ubc9. J. Biol. Chem. *276*, 35368–35374.

140. Tiefenbach, J., Novac, N., Ducasse, M., Eck, M., Melchior, F., and Heinzel, T. (2006). SUMOylation of the Corepressor N-CoR Modulates Its Capacity to Repress Transcription. Mol. Biol. Cell *17*, 1643–1651.

141. Titus, M.A., Schell, M.J., Lih, F.B., Tomer, K.B., and Mohler, J.L. (2005). Testosterone and Dihydrotestosterone Tissue Levels in Recurrent Prostate Cancer. Clin. Cancer Res. *11*, 4653–4657.

142. Tolis, G., Ackman, D., Stellos, A., Mehta, A., Labrie, F., Fazekas, A.T., Comaru-Schally, A.M., and Schally, A.V. (1982). Tumor growth inhibition in patients with prostatic carcinoma treated with luteinizing hormone-releasing hormone agonists. Proc. Natl. Acad. Sci. U. S. A. *79*, 1658–1662.

143. Treuter, E., and Venteclef, N. (2011). Transcriptional control of metabolic and inflammatory pathways by nuclear receptor SUMOylation. Biochim. Biophys. Acta BBA - Mol. Basis Dis. *1812*, 909–918.

144. van der Veen, A.G., and Ploegh, H.L. (2012). Ubiquitin-like proteins. Annu. Rev. Biochem. *81*, 323–357.

145. Veldscholte, J., Ris-Stalpers, C., Kuiper, G.G.J.M., Jenster, G., Berrevoets, C., Claassen, E., van Rooij, H.C.J., Trapman, J., Brinkmann, A.O., and Mulder, E. (1990). A mutation in the ligand binding domain of the androgen receptor of human INCaP cells affects steroid binding characteristics and response to anti-androgens. Biochem. Biophys. Res. Commun. *173*, 534–540.

146. Visakorpi, T., Hyytinen, E., Koivisto, P., Tanner, M., Keinänen, R., Palmberg, C., Palotie, A., Tammela, T., Isola, J., and Kallioniemi, O.-P. (1995). In vivo amplification of the androgen receptor gene and progression of human prostate cancer. Nat. Genet. *9*, 401–406.

147. Vucic, D., Dixit, V.M., and Wertz, I.E. (2011). Ubiquitylation in apoptosis: a post-translational modification at the edge of life and death. Nat. Rev. Mol. Cell Biol. *12*, 439–452.

148. Vyas, R., Kumar, R., Clermont, F., Helfricht, A., Kalev, P., Sotiropoulou, P., Hendriks, I.A., Radaelli, E., Hochepied, T., Blanpain, C., et al. (2013). RNF4 is required for DNA double-strand break repair in vivo. Cell Death Differ. *20*, 490–502.

149. Walsh, C.T., Garneau-Tsodikova, S., and Gatto, G.J. (2005). Protein Posttranslational Modifications: The Chemistry of Proteome Diversifications. Angew. Chem. Int. Ed. *44*, 7342–7372.

150. Wang, L., Wansleeben, C., Zhao, S., Miao, P., Paschen, W., and Yang, W. (2014). SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. EMBO Rep. *15*, 878–885.

151. Wang, Q., Xia, N., Li, T., Xu, Y., Zou, Y., Zuo, Y., Fan, Q., Bawa-Khalfe, T., Yeh, E.T.H., and Cheng, J. (2013). SUMO-specific protease 1 promotes prostate cancer progression and metastasis. Oncogene *32*, 2493–2498.

152. Watanabe, T.K., Fujiwara, T., Kawai, A., Shimizu, F., Takami, S., Hirano, H., Okuno, S., Ozaki, K., Takeda, S., Shimada, Y., et al. (1996). Cloning, expression, and mapping of UBE2I, a novel gene encoding a human homologue of yeast ubiquitin-conjugating enzymes which are critical for regulating the cell cycle. Cytogenet. Genome Res. *72*, 86–89.

153. Waters, T.R., Gallinari, P., Jiricny, J., and Swann, P.F. (1999). Human Thymine DNA Glycosylase Binds to Apurinic Sites in DNA but Is Displaced by Human Apurinic Endonuclease 1. J. Biol. Chem. *274*, 67–74.

154. Watson, P.A., Arora, V.K., and Sawyers, C.L. (2015). Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. Nat. Rev. Cancer *15*, 701–711.

155. Wei, W., Yang, P., Pang, J., Zhang, S., Wang, Y., Wang, M.-H., Dong, Z., She, J.-X., and Wang, C.-Y. (2008). A stress-dependent SUMO4 sumoylation of its substrate proteins. Biochem. Biophys. Res. Commun. *375*, 454–459.

156. Wenzel, D.M., Stoll, K.E., and Klevit, R.E. (2011). E2s: structurally economical and functionally replete. Biochem. J. *433*, 31–42.

157. Wijngaart, D.J. van de, Molier, M., Lusher, S.J., Hersmus, R., Jenster, G., Trapman, J., and Dubbink, H.J. (2010). Systematic Structure-Function Analysis of Androgen Receptor Leu701 Mutants Explains the Properties of the Prostate Cancer Mutant L701H. J. Biol. Chem. *285*, 5097–5105.

158. Wilson, V.G. (2009). Introduction to Sumoylation. In SUMO Regulation of Cellular Processes, V.G. Wilson, ed. (Springer Netherlands), pp. 1–10.

159. Wilson, V.G. (2017). Introduction to Sumoylation. In SUMO Regulation of Cellular Processes, V.G. Wilson, ed. (Springer International Publishing), pp. 1–12.

160. Xu, P., Duong, D.M., Seyfried, N.T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D., and Peng, J. (2009). Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. Cell *137*, 133–145.

161. Yang, S.-H., Galanis, A., Witty, J., and Sharrocks, A.D. (2006). An extended consensus motif enhances the specificity of substrate modification by SUMO. EMBO J. *25*, 5083–5093.

162. Zhang, F.-P., Mikkonen, L., Toppari, J., Palvimo, J.J., Thesleff, I., and Jänne, O.A. (2008). Sumo-1 Function Is Dispensable in Normal Mouse Development. Mol. Cell. Biol. *28*, 5381–5390.

163. Zhang, H., Saitoh, H., and Matunis, M.J. (2002). Enzymes of the SUMO Modification Pathway Localize to Filaments of the Nuclear Pore Complex. Mol. Cell. Biol. *22*, 6498–6508.

164. Zhang, L., Vogel, W.K., Liu, X., Topark-Ngarm, A., Arbogast, B.L., Maier, C.S., Filtz, T.M., and Leid, M. (2012). Coordinated Regulation of Transcription Factor Bcl11b Activity in Thymocytes by the Mitogen-activated Protein Kinase (MAPK) Pathways and Protein Sumoylation. J. Biol. Chem. *287*, 26971–26988.

165. Zhao, B., Bhuripanyo, K., Schneider, J., Zhang, K., Schindelin, H., Boone, D., and Yin, J. (2012). Specificity of the E1-E2-E3 Enzymatic Cascade for Ubiquitin C-Terminal Sequences Identified by Phage Display. ACS Chem. Biol. *7*, 2027–2035.

166. Zhao, X.-Y., Malloy, P.J., Krishnan, A.V., Swami, S., Navone, N.M., Peehl, D.M., and Feldman, D. (2000). Glucocorticoids can promote androgen-independent growth of prostate cancer cells through a mutated androgen receptor. Nat. Med. *6*, 703–706.

167. Zhou, Y.-F., Liao, S.-S., Luo, Y.-Y., Tang, J.-G., Wang, J.-L., Lei, L.-F., Chi, J.-W., Du, J., Jiang, H., Xia, K., et al. (2013). SUMO-1 Modification on K166 of PolyQ-Expanded aTaxin-3 Strengthens Its Stability and Increases Its Cytotoxicity. PLOS ONE *8*, e54214.

Chapter 2

Generation of a robust FRET-based assay for high throughput screening and characterization of SENP enzymes

1. Introduction

SUMO and the components involved in SUMOylation and deSUMOylation process are frequent targets of study because of their important roles in most cellular processes. The embryonic lethality in mice of E2 Ubc9, SUMO2, SENP1, and SENP2 knockouts supports the notion that dynamic SUMO modification is essential for exquisitely modulating protein activity in development and beyond (Nacerddine et al., 2005; Wang et al., 2014; Sharma et al., 2013; Kang et al., 2010). These characteristics and the novelty of the SUMO system with its short two decades since discovery mean that several groups have generated an array of tools attempting to study the system.

Highlighted in chapter 1, the family of SUMO proteases, known as SENPs, have been of particular interest due to their role in numerous cancers (Eifler and Vertegaal,

‡ Many of the experiments in this dissertation were carried out with participation of multiple individuals. For each relevant chapter, the contributions are indicated in a footnote in the first page of each chapter. Carrie M. Johnson performed the assay condition optimization and kinetic characterization. Carrie M. Johnson performed the substrate characterization and assay performance in collaboration with Dr. Jorge Iñiguez-Lluhí. The substrate design expression and purification were carried out by Dr. Jorge Iñiguez-Lluhí.

2015). Because of the similarities to ubiquitin, many of the tools used to study deubiquitinating enzymes were adapted for the SUMO system. One of the earliest tools uses a hemagglutinin (HA) tagged SUMO with a vinyl sulfone (VS) moiety attached to the Gly-Gly C-terminal SUMO tail. When in proper proximity and confirmation, the VS group forms a covalent adduct with the catalytic cysteine of the SENP enzyme. This irreversible inhibitor can be used to probe the different SENPs for their preference for the SUMO isoforms. SUMO-AMC is a fluorogenic substrate where the C-terminal glycine residue of SUMO is conjugated to 7-amino-4-methylcoumarin (AMC). When still conjugated to SUMO, the intrinsic fluorescence of AMC is quenched. SENP enzymatic activity cleaves AMC from SUMO, releasing AMC and increasing the fluorescence of the system at 460 nm and allowing measurement of SENP activity using steady state kinetics (Kolli et al., 2010; Madu and Chen, 2001).

Like SUMO-AMC, the C/EBP homologous protein (CHOP)-reporter platform links SUMO and another chemical moiety, this time an enzyme, phospholipase A_2 (PLA_2). Attached at its amino-terminus to the C-terminal glycine residue of SUMO, PLA_2 is an enzyme that only functions when its N-terminus is free (Nicholson et al., 2008). Only upon cleavage and release from SUMO by SENP protease activity does the PLA_2 become active. Acting as a reporter enzyme, PLA_2 is able to cleave its substrate NBD C_6-HPC and generate the fluorescent product, NBD, to generate a detectable, quantitative fluorescent readout of SENP activity (Leach et al., 2009). Alternatively, because PLA_2 is an enzyme of approximately 14 kDa, a gel based approach monitoring the shift in molecular weight of the starting SUMO-CHOP construct to free SUMO and free PLA_2 can be used instead of

fluorescence. In this format, no PLA_2 substrate is needed but the assay has lower throughput.

In a different kind of gel-based approach, SUMO-conjugated Ran GTPase-activating protein (RanGAP) is used to study SENP isopeptidase activity. Unlike the linear SUMO-based tools that rely on SENP endopeptidase activity, this system requires more effort to generate the substrate by combining SUMO, the SUMO E1 and E2 enzymes, and the SUMO target protein, RanGAP. Once SUMOylation is achieved, the SUMO-RanGAP substrate can be incubated with a SENP enzyme and the reaction can be killed via addition of SDS sample buffer at different time points or in a dose responsive manner and run on a gel for general quantitation.

In a bioluminescent approach to study SENP activity, a commercially available SENP substrate of carboxybenzyl-Arg-Leu-Arg-Gly-Gly-luciferin is used to assess protease activity. Alternatively, an amino-luciferin (AML) analog of this substrate provides further sensitivity from the original luciferin substrate (Orcutt et al., 2012). The SENP enzyme recognizes the Gly-Gly sequence and cleaves the peptide bond after the second glycine residue to release free luciferin. Luciferin is a substrate for the enzyme luciferase, and generating the bioluminescent substrate oxyluciferin. Though this reporter enzyme system, the bioluminescence approach has the advantage of improved sensitivity and wider dynamic range over substrates with a quenched fluorophore conjugated to the peptide C-terminus like SUMO-AMC (Leippe et al., 2011). In testing for inhibitors, this approach also avoids issues related to fluorescence interference by test compounds. However, as a small peptide substrate, this system does not allow for differentiation among the SUMO isoforms. It also misses the important contributions the globular portion of

SUMO has in contact with SENPs and yields kinetic parameters orders of magnitude different from those determined using full length SUMO (Chen et al., 2014).

Fluorescence resonance energy transfer (FRET) assays have been adapted for the SUMO/SENP system. The FRET system for studying the SENP enzymes has experienced a series of improvements to enhance FRET efficiency and quantitation of SENP kinetics. The substrate has been optimized to generate better FRET efficiency and quantum yield and the cross contamination of fluorescent signal had been accounted for, but shortcomings remain (Jiang et al., 2013, 2014; Liu and Liao, 2013; Liu et al., 2012, 2015, Martin et al., 2007, 2008). Despite the availability of several assay formats for SENP1, for the most part, the assays have been used in a semi-quantitative manner without conversion to actual substrate or product concentration values and without applying a rigorous enzyme kinetic analysis. Of key significance, SENP enzymes utilize a large exosite for recognition away from the active site and are thus expected to be subject to significant product inhibition. However, no satisfactory consideration has been given to this process in the existing literature, which can severely influence estimation of kinetic parameters and mechanism of action of inhibitors. Given the interest in the potential therapeutic applications of SENP based therapeutics, a successful drug discovery and development program depends critically on a robust and well-validated assay pipeline. Robust, scalable and high throughput ready assays are essential for hit identification and validation whereas quantitative detailed kinetic characterization assays are indispensable for defining the properties of the enzyme and the mechanism of action of inhibitors. Recognizing these needs and the shortcomings of existing platforms, our group has generated and characterized in detail a set of evolved FRET-based SENP substrates that can be used to

monitor the peptidase activity of SENP and can be readily obtained in high yield. We have used them to develop robust, miniaturized and facile ratiometric assays for high-throughput screening and selectivity assessment. By leveraging our detailed characterization of the substrates, we have also developed a wide dynamic range quantitative assay as well as numerical integration and global fitting algorithms for data analysis. This combination allows for simultaneous analysis of entire progress curves across a broad range of substrate, product and inhibitor concentrations. By fitting to kinetic models that explicitly account for product inhibition, it is possible to obtain precise determinations of kinetic parameters and assign specific modes of inhibition to active compounds. Furthermore, analysis of SENP1 mutants and the effects of ionic strength has generated insight into important residues for SENP1 substrate recognition and activity, the crucial role of electrostatic interactions, and key differences between SENP1 and its closest paralog SENP2. These features can be leveraged in efforts to enhance selectivity of inhibitors during drug development.

2. Materials and Methods

FRET substrate construction, expression, and purification. SFCypet-SUMO1-SFYpet and CyPet-SUMO2-SFYPet were constructed in pET15b plasmids with expression under control of the lac promoter. A tobacco etch virus protease cleavage sequence was introduced at the C-terminus followed by a 10x-His tag. Plasmids were obtained from Dr. Joshua Plotkin to contain a 5' untranslated region for high yield *E. coli* expression determined using experimental observation (Kudla et al., 2009). Additional mutations to

the base CFP and YFP coding sequences were made to generate a FRET-optimized pair for improved energy transfer, sensitivity, dynamic range, generating Cypet and Ypet (Nguyen and Daugherty, 2005). Lastly, a final set of mutations were introduced to generate a superfolder (SF) version of each fluorescent protein to limit aggregation and misfolding (Pédelacq et al., 2006). In addition, a similar construct harboring SUMO 2 was generated. Expression was carried out in Rosetta 2 (DE3) *E. coli* cells expressing rare tRNAs. Large-scale production was obtained in a single 20L fermentor run. Initial lysis and extract generation was carried out as described below for the catalytic domains of SENP1 and SENP2. Purification involved sequential Ni-chelate chromatography (Ni-NTA), desalting (G-50), anion exchange (SourceQ) and size exclusion chromatography (Superdex 200).

SENP1 and SENP2 catalytic domain construction, expression, and purification. N-terminal His-tagged SENP1 (AA 418-C) and SENP2 (AA 363-C) catalytic domains were constructed in pHT2 plasmids under control of the lac promoter. Plasmids were grown under carbenicillin selection and transformed in to Rosetta 2 (DE3) *E. coli* cells. Cells were grown at 37 °C until OD_{600} 1.1, upon which the temperature was dropped to 18 °C and isopropyl β-D-1-thiogalactopyranoside was added to 0.2 mM final. Cells were grown for another 20 hours. After pelleting cells, they were re-suspended in resuspension buffer (10 mM Tris pH8.0, 500 mM NaCl, 5% glycerol, 2 mM imidazole, 1 mM BME, 0.5 mM PMSF, 0.5 μL benzonase) and lysed via French press for three cycles. Cell lysate was centrifuged at 40,000 x g for 45 min at 4°C. Protein was purified from the lysate using Ni-NTA affinity chromatography and stored in 250 mM NaCl, 50 mM Tris pH 7.5, and 1 mM dithiothreitol.

Generation of substrate samples with defined cleavage fractions. Samples of fully digested SUMO1 substrate were generated by incubating SUMO1 substrate (21 μM) with SENP 1CD (21.5 nM) for 6 hours at room temperature. Enzyme activity was inactivated via addition of N-ethylmaleimide (5 mM) and incubated for 30 minutes. Unreacted NEM was quenched with Dithiothreitol (20 mM) for 10 minutes. A sample of intact substrate containing the same components and incubated in parallel was obtained in the same manner but by altering the order of addition so that the SENP1 CD was inactivated before mixing with the substrate. The reactant and product sources were then mixed in different ratios to generate stable samples containing fractions of cleaved substrate of 0, 0.2, 0.4, 0.6, 0.8 and 1.

Fluorescence spectra. Samples of both intact and fully digested SUMO1 substrate as well as samples with intermediate cleavage fractions of the substrate were prepared as indicated above. Fluorescence spectra were obtained in a HORIBA fluoromax3 fluorimeter at room temperature using 2 nm excitation and emission slit widths. Scans were obtained using a rectangular microcuvette with 0.1 cm excitation and 1 cm emission pathlengths. No correction for primary or secondary inner filter effects was necessary since calculated absorbance based on absorption spectra of the samples remained <0.01 across both the excitation and emission pathlengths at the wavelengths probed. Emission scans from 450 to 650 nm were obtained at 1 nm resolution and 0.1s integration time at excitation wavelengths from 300 to 600 nm at 10 nm intervals or at the wavelength indicated in the figures. Emission data were corrected for lamp intensity at the specific excitation

wavelength in real time. Data from triplicate scans were averaged and the values from corresponding blank samples were subtracted. Raman as well as first and second order Rayleigh scatter signals remaining after blank subtraction were suppressed by applying a high band pass filter and smoothing in order to generate the Excitation/emission matrix data in Figure 2-1C.

Melting temperature determination. Thermal shift experiments were carried out using a Thermofluor 384 ELS system (Johnson & Johnson). Protein unfolding was examined by monitoring the fluorescence of ANS (1-anilinonaphthalene-8-sulfonic acid) by increasing the temperature from 25° to 60°C. All samples were prepared in quadruplicate. Final protein sample concentration was 0.2 mg/mL, and 0.1 mM ANS. To limit evaporation, the samples were covered with 1.5 μL of silicon oil. Buffers or salts tested for stability were added as a 2X stock to the protein and dye mixture. The reaction mixture was allowed to equilibrate for 30 min at room temperature in the dark before beginning the denaturation experiments. Melting temperatures were obtained as the inflection points of the fluorescence signal as a function of temperature.

***In Vitro* ratiometric FRET assay**. For 96-well assays, reactions were 100 μL final and performed in high or low salt buffer solution (20 mM TrisCl pH 8.0, 250 μM Na·EDTA, 20 mM NaCl or 100 mM NaCl, 10 mM β-Mercaptoethanol, 0.1 mg mL^{-1} bovine serum albumin, 0.01% NP-40). Enzyme concentration ranged from 75-200 pM depending on the number of plates and frequency of reads. Substrate concentration was 100 nM. Time delays between addition of enzyme to start reaction for each well and the first read by the

instrument were accounted for. A Spectramax M5 plate reader was used for data capture with excitation at 405 nm and emission measured at 475 and 525 nm. To adapt the assay to 384 well formats, a scaling to a 20 μL final volume was used and data were acquired in an Envision plate reader with CFP and YFP filters.

Kinetic analysis SENP assay. Assays were carried out in a 96 well format essentially as described for the ratiometric assay at concentrations of substrate ranging from 3 nM to 2 μM in the absence or presence of various concentrations of inhibitor (purified free SUMO1) ranging from 7 nM to 200 nM. The amount of enzyme in each well was adjusted based on kinetic modeling (see below) in order to achieve progress curves of comparable time trajectories. This approach allows for monitoring of a wide range of substrate and inhibitor combinations over the same time frame and after exposure to the same illumination. This permits comparisons without concerns about differential photobleaching or other time-dependent effects. Samples of intact and completely digested substrate (obtained by incubation with a large excess of enzyme) at each of the concentrations of substrate tested were monitored in parallel.

Data analysis and model fitting. Data for kinetic characterization obtained as described above were processed as follows: Raw channel 2 data was blank subtracted and corrected for non-linearity with respect to concentration by applying a second-degree polynomial correction factor derived from analysis of the standard curve data. The correction factor values increase with concentration but did not exceed ~15% at the highest substrate concentrations used (~4 μM). Corrected channel 2 data were converted to fraction of

substrate cleaved (F) assuming that the signal observed is a linear combination of the contributions of the substrate and product using the following formula:

$$F(t) = \frac{Ch_2(t_0) - Ch_2(t)}{Ch_2(t_0) - Ch_2(t_\infty)}$$

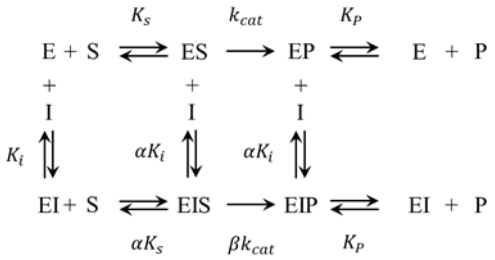
with $Ch2(t_0)$ and $Ch2(t_\infty)$ corresponding to the fluorescence values of the intact and completely digested substrate obtained in parallel at each substrate concentration. The amount of product generated as a function of time at each initial substrate concentration was obtained by multiplying $F(t)$ by the corresponding substrate concentration. Fitting to obtain kinetic parameters was carried out by iterative minimization of the global normalized sum of square differences between all observed and predicted values across all combinations of conditions. Predicted values were generated at each iteration by numerical integration (fourth order Runge-Kutta) of the rate equation for a rapid equilibrium mixed type inhibition model with product inhibition (Segel, 1975).

Equation 1:

$$v = \frac{E_t k_{cat} \left(\frac{[S]}{K_s} + \frac{\beta [S][I]}{\alpha K_s K_i} \right)}{1 + \frac{[S]}{K_s} + \frac{[P]}{K_p} + \frac{[I]}{K_i} + \frac{[S][I]}{\alpha K_s K_i}}$$

Corresponding to the following reaction scheme:

Scheme 1:



where K_s and K_p correspond to substrate and product dissociation constants and k_{cat} as the catalytic rate constant. This model accommodates mixed types of inhibition. If substrate and inhibitor binding is mutually exclusive (competitive mechanism), $\alpha \gg 1$. In pure noncompetitive inhibition, $\beta \ll 1$ and $a \sim 1$ whereas in pure uncompetitive inhibition, $\beta = 0$ and $a \ll 1$.

Numerical integration and minimization were carried out in Microsoft Excel using a custom spreadsheet and the GRG nonlinear solver with forward derivatives and a convergence threshold of 10^{-4}. Standard deviations for the parameters were obtained from independent fitting of >250 bootstrapped datasets derived from the experimental data.

3. Results

FRET Substrate Generation

The FRET substrate was designed with the pair-optimized Cypet at the N-terminus of SUMO with a small peptide linker between the fluorophore and SUMO. Following the SUMO maturation recognition site at the C-terminus is a peptide linker and the other member of the FRET pair, Ypet (Figure 2-1A). Using the superfolder (SF) versions of the fluorophores allowed for the production of highly purified protein in substantial quantities (data in possession of Dr. Jorge Iniguez) (Figure2-1B, lane 1). In this set-up, the linkers allow the SFCypet-SFYpet pair to be in close proximity for efficient energy transfer while allowing access by SENPs to the SUMO Gly-Gly cleavage site to generate the SFCypet-SUMO1 and the SFYpet products (Figure 2-1B, lane 6).

FRET Substrate Spectral Characterization

Purified SFCypet-SUMO1-SFYpet substrate fluorescent properties were assessed to determine the optimal wavelength for excitation as well as the wavelengths for detecting SFCypet and SFYpet signal for both intact substrate and cleaved products. Analysis of samples of intact substrate and fully cleaved product species for fluorescent analysis were confirmed for their expected pattern of species via SDS-PAGE (Figure 2-1B lanes 1 and 6). Fluorescent spectra on intact substrate samples show that excitation of substrate at 405 nm generates a prominent fluorescence signal at emission wavelengths of 525 nm, the signal dominated by SFYpet emissions (Figure 2-1C, left). The emission signal at 475 nm, which pertains to SFCypet, is relatively weak under these conditions. In contrast, excitation at ~500 nm directly excites the SFYpet, generating an emission signal at 525 nm and no signal at 475 nm. In the fully cleaved product, similar SFCypet excitation at 405 nm generates an emission spectrum with a drastically reduced signal at 525 nm and an enhanced signal at 475 nm (Figure 2-1C, right). By using excitation at 405 nm, the direct excitation of SFYpet is limited (lower 525 emission) while still achieving SFCypet excitation. The dynamic range in signal between intact substrate and cleaved product is much larger for the signal at 525 nm (approximately 10-fold) than 475 nm (Figure 2-1D). The remaining 525 nm emission signal in fully cleaved product is a result of minimal direct SFCypet excitation. Based on these results, to assess SENP protease activity in our FRET-based assays an excitation wavelength of 405 nm was used with emission measured at 475 nm (channel 1) and 525 nm (channel 2). Protease activity can be monitored via the loss in channel 2 signal and concurrent increase in channel 1 signal due to loss of FRET.

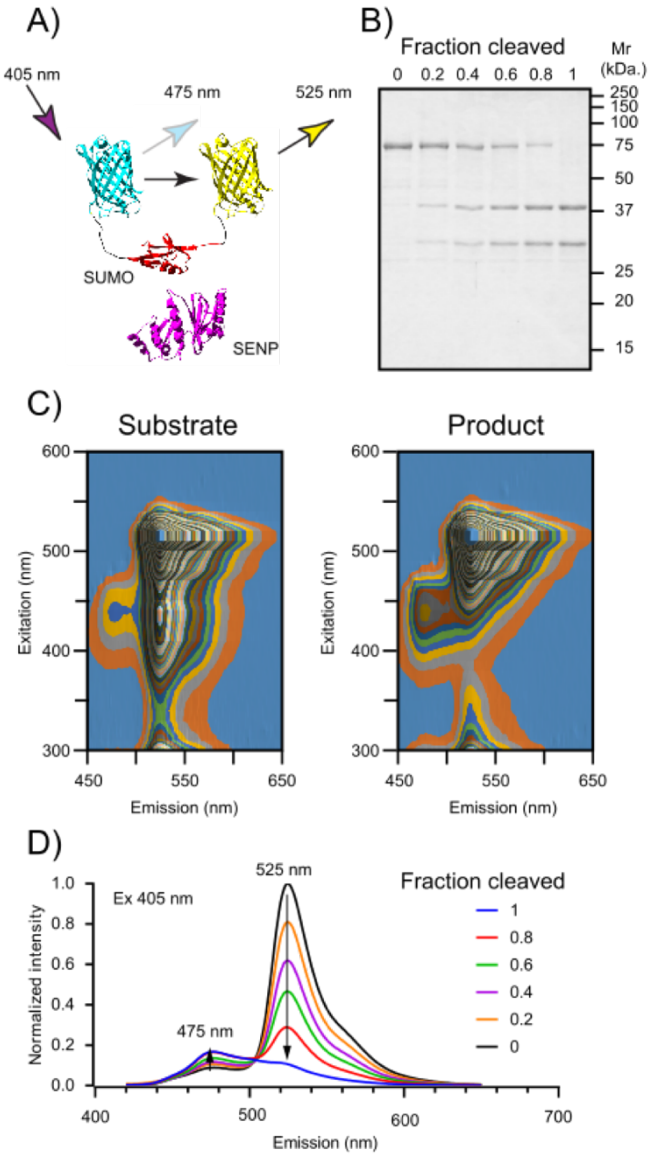


Figure 2-1 Substrate Characterization.
A) Depiction of FRET substrate organization showing excitation of SFCypet at 405 nm with transfer of the energy in intact substrate to SFYpet, leading to fluorescence emission at 525 nm. After protease activity on the substrate by SENP, SFYpet is no longer in proximity to SFCypet and emission at 475 nm is increased. B) SDS-PAGE of different ratios of intact FRET substrate (75 kDa) and the fully cleaved product (42 kDa and 33 kDa). C) Fluorescence spectra of intact substrate (left) and cleaved/mature product (right). D) Fluorescence intensity after excitation at 405 nm for the six ratios of cleaved product to intact substrate.

Ratiometric Assay Characterization

With the substrate conditions optimized, the behavior of the assay and its ability to measure SENP protease activity in real time was assessed. As the reaction progresses, the channel 2 signal decreases while the channel 1 signal increases, but not linearly due to significant product inhibition (Figure 2-2A). The ratiometric change that can be plotted as a function of time to assess enzyme activity (Figure 2-2B). As seen in the initial substrate characterization, the magnitude of change in channel 2 signal over time is much larger than

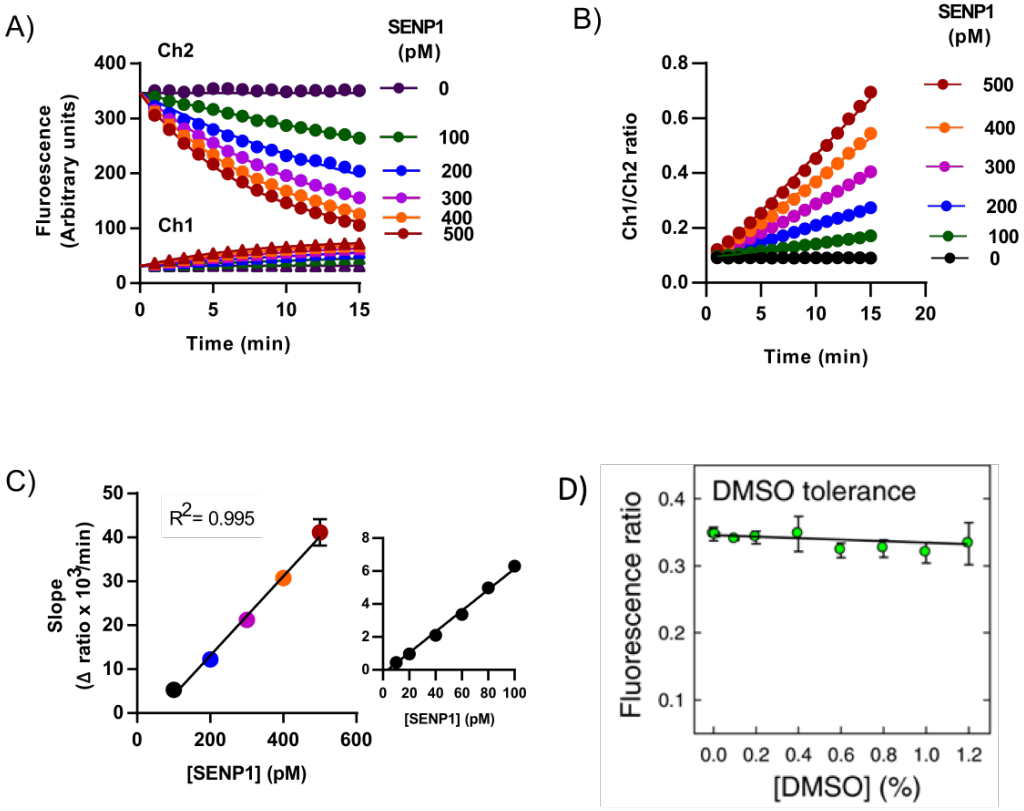


Figure 2-2 FRET assay performance.
A) Change in channel 1(475 nm) and channel 2 (525 nm) fluorescence signal over time at different concentrations of SENP1 enzyme present. B) Plots of the ratios of channel 1/channel 2 using the data from (A). C) The slopes of the ratios over time from panel (B) plotted as a function of enzyme concentration. D) The effects of DMSO on FRET ratios showing that the assay is tolerant to high levels of DMSO.

the magnitude for channel 1. The ratiometric readout functions as an index of enzyme activity and the slopes of these ratios as a function of enzyme concentration display a linear relationship (Figure 2-2C). In addition to enabling assessment of SENP activity across a large range of enzyme concentrations and over a long period of time, the assay is also quite tolerant of the presence of DMSO (Figure 2-2D). The tolerance to DMSO is important for use of the assay to test potential small molecule inhibitors, as often these small molecules are stored and dissolved in it. With a goal of discovering small molecule SENP1 inhibitors, this DMSO tolerance is essential for a robust assay for future high-throughput screening. The ratiometric approach allows for slight variations in substrate concentration to be inconsequential. This allows for adaptation of the assay to miniaturization to 384 and 1536 well formats to increase the throughput for studying SENPs. Under the ratiometric approach the assay can be utilized in endpoint, dose response, and time course assays to characterize inhibitors for intrinsic fluorescence, selectivity to SENP isoforms, and to develop structure activity relationships among families of small molecule inhibitors.

SENP enzyme condition optimization in the FRET assay

The protease activity of the catalytic domains of SENP1 and SENP1's closest paralog, SENP2, were evaluated under different pH and salt conditions. The salt sensitivity of SENP1 relative to SENP2 differ dramatically (Figure 2-3A). As the concentration of salt increases from 20 mM to 100 mM, the maximum catalytic activity of SENP1 decreases throughout while SENP2 approaches maximal activity. Whether NaCl, LiCl, or KCl, the SENP salt sensitivity curves behaved the same way (data not shown). In contrast to the differing behaviors between the two enzymes concerning salt, their changes in stability and

activity at differing pH values were more congruent. The lack of destabilization of the enzymes, suggesting that the effects of ionic strength on SENP activity may reflect effects

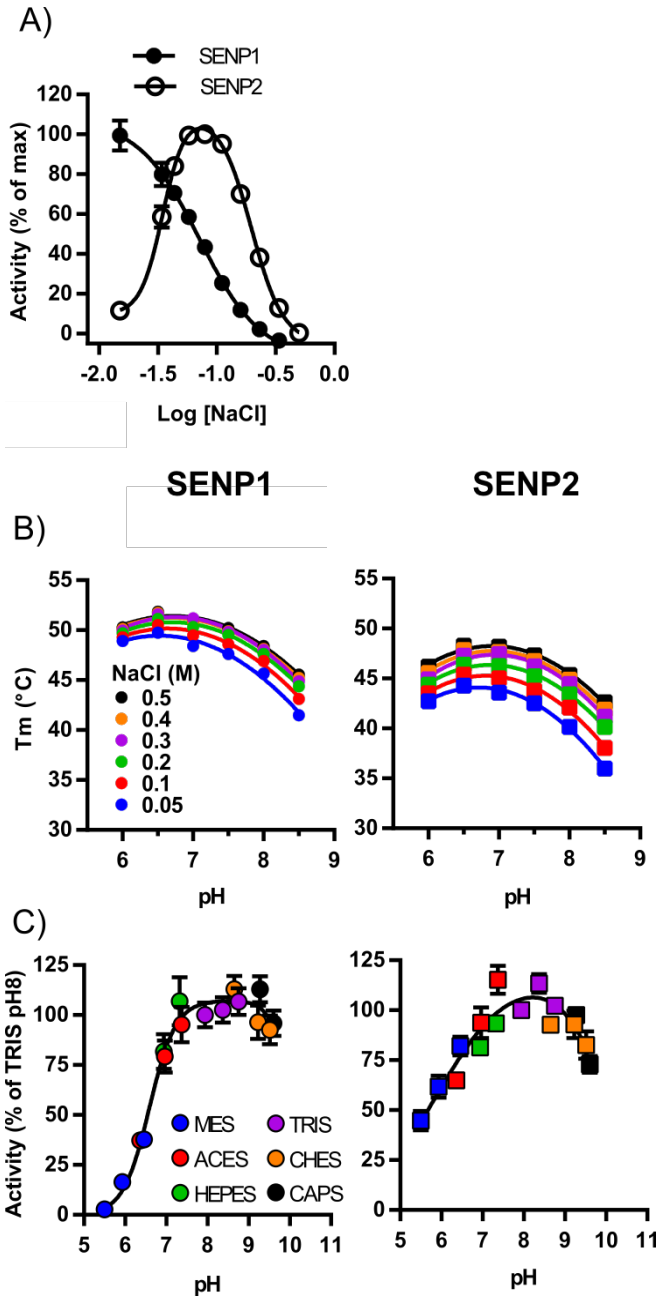


Figure 2-3 SENP condition optimization for FRET assay.
A) SENP1 and SENP2 activity were monitored via the ratiometric FRET assay under different salt conditions. In addition to NaCl (shown), LiCl and KCl were also tested with similar results in salt sensitivity. B) Thermofluor-determined melting points of SENP1 and SENP2 at different pH and salt conditions. C) Using the ratiometric FRET assay, SENP activity determined as a function of pH.

on substrate interactions more than the intrinsic properties of the SENPs in isolation. Both enzymes show the most thermal stability at a pH near 6.5 to 7 and show optimal catalytic activity against the FRET substrate at a pH between 8 and 9 (Figure 2-3B and C). From these results, buffer conditions of pH 8 were used for all SENP assays and 20 mM NaCl was used for SENP1 while 100 mM was used for SENP2 unless otherwise noted.

Kinetic Characterization of SENP enzymes

While the ratiometric analysis is a robust tool for screening conditions and potential small molecule inhibitors, at low substrate concentrations, the ratio or enzyme rate is susceptible to small, noise-related changes in channel 1 signal because of its small dynamic range and signal to noise ratio. Yet low substrate concentrations are needed to obtain kinetic constants for SENPs and their SUMO processing. Because of the high affinity interaction between the two and the need to test at concentrations above and below these binding constants, we devised a quantitative approach based on the channel 2 signal alone instead of the ratio of channel 1 to channel 2.

Since the goals of this thesis are centered on the discovery and development of small molecule inhibitors of SENP1 and much of the binding affinity between SUMO1 and SENP1 is driven by interactions with the globular β-grasp domain, significant product inhibition is expected and mature free SUMO1 would be expected to behave as a competitive inhibitor. We therefore tested the effects of purified mature free SUMO1 as an inhibitor. Using a matrix of conditions where the concentrations of SENP1 WT catalytic domain, SFCypet-SUMO1-SFYpet, and free SUMO1 were varied to maximize the coverage of data points at critical points related to substrate and product binding affinities

(Figure 2-4A). Channel 2 fluorescent values were converted to product concentration over time as described under materials and methods.

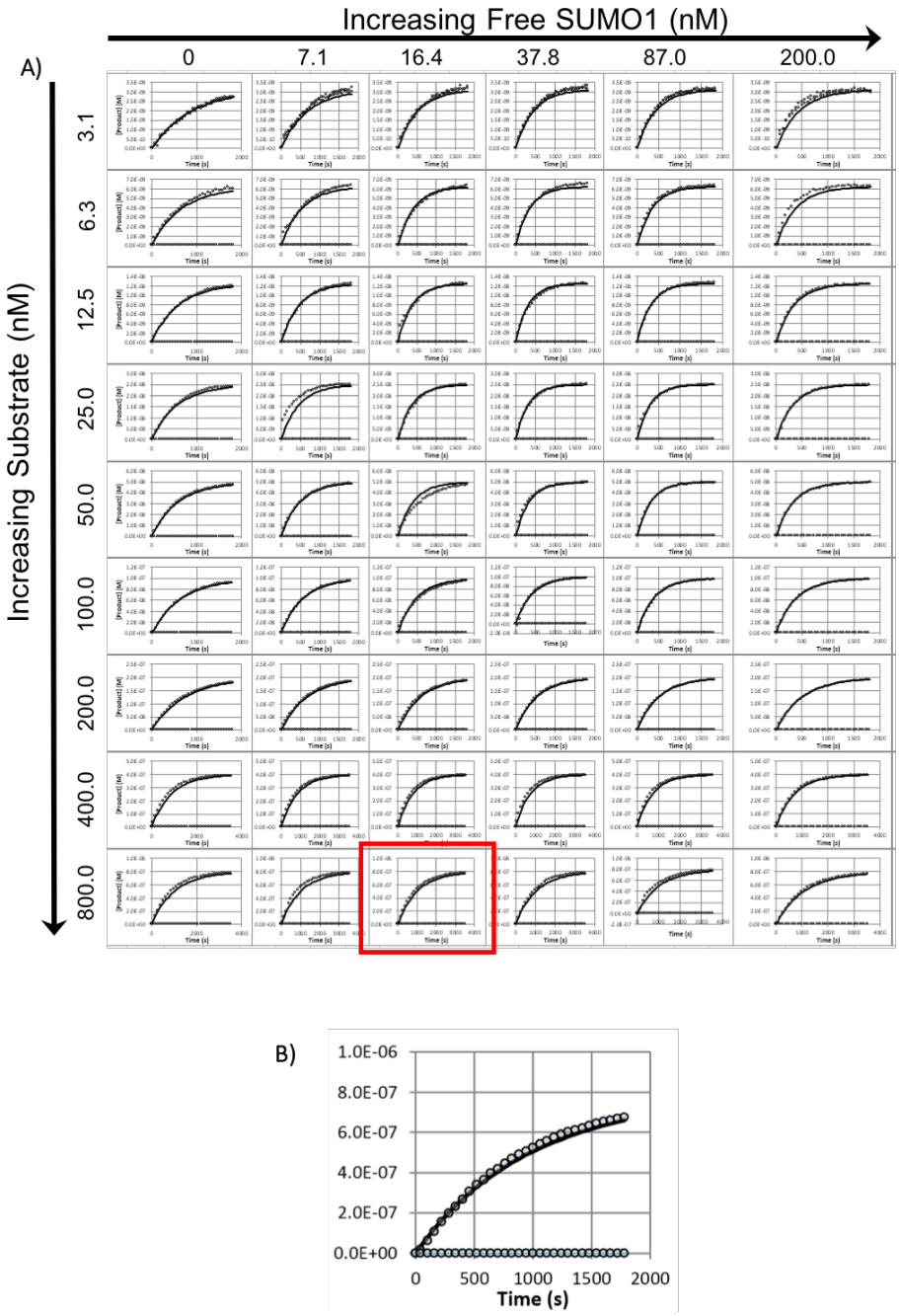


Figure 2-4 Kinetic characterization of SENP1 with free SUMO1 inhibitor.
A) Representative progress curves for the matrix of substrate and mature free SUMO1 inhibitor conditions used to determine SENP1 kinetic constants and the inhibitory constant of mature free SUMO1. Fits obtained are derived from simultaneous global fitting of the model to all the progress curves. B) Representative curve from (A) enlarged to show individual progress curve fitting.

The results of global fitting revealed in an unbiased manner that mature free SUMO1 is a competitive inhibitor as anticipated since when all parameters are allowed to float, minimization yielded a value of $\alpha = 23.1$ which under the range of concentrations of substrate analyzed would imply essentially mutually exclusive binding between substrate and inhibitor. Further analysis carried out with a purely competitive scheme (fixing α to a very large number and β to 0) yielded values for the kinetic parameters k_{cat}, K_s, and K_p for SENP1, where K_s is the binding affinity for substrate and K_p is the binding affinity for the product and K_i is the affinity for free SUMO1 (Figure 2-5A). The SENP1 K_s and K_p values are nearly the same and the K_i value for free SUMO1 is in the same order of magnitude. These values clearly highlight for the first time in a rigorous manner that SENP1 is subject to extensive product inhibition. The similar affinities for free SUMO1 and the Cypet SUMO1 product is not surprising given the lack of binding affinity from the C-terminal tail residues near the SUMO processing site based on the crystal structure of the complex and the retention of nearly all residues contributing to the binding affinity present in the processed or mature forms of substrate. This also indicates that the fluorophores on the FRET substrate are simply serving as a chemical tool to monitor SENP activity and have little effect on substrate or product binding.

With the ability to determine SENP kinetic constants using the channel 2 analysis and considering the differential salt sensitivity of SENP1 and SENP2, the two enzymes were tested under their alternative salt conditions. Essentially, SENP1 was tested at SENP2's ideal salt concentration and vice-a-versa. As seen in Figure 2-5B, the k_{cat} value for SENP1 does not change when salt concentration is increased, but its affinity for substrate and cleaved product are weakened substantially. In contrast, for SENP2 at low

salt concentration, its k_{cat} is halved and affinity for substrate and product is weakened by more than three-fold and more than eight-fold respectively. Treating substrate affinity as an apparent K_m, the catalytic efficiency of SENP2 (k_{cat}/K_s) drops from around 40×10^6 $M^{-1}s^{-1}$ at high salt to approximately 5.5×10^6 $M^{-1}s^{-1}$ at low salt. Similar analysis for SENP1

A)

[NaCl]	20 mM
Enzyme	SENP1 WT
Inhibitor	Free SUMO1
k_{cat} (s^{-1})	0.85 (0.13)
K_s (nM)	6.87 (0.13)
K_p (nM)	5.96 (0.12)
Ki (nM)	17.3 (0.33)

B)

	SENP1		SENP2	
[NaCl] mM	20	100	20	100
k_{cat} (s^{-1})	1.12 (0.14)	1.04 (0.06)	309 (0.004)	0.69 (0.17)
K_s (nM)	10.4 (2.9)	384 (17.3)	55.7 (0.9)	17.2 (3.5)
K_p (nM)	6.4 (2.3)	201 (7.9)	100 (3.7)	12.3 (3.1)

C)

	SENP1 (20 mM NaCl)				SENP2 (100 mM NaCl)
SENP type	WT	K455G	K515G	K455,515G	WT
k_{cat} (s^{-1})	1.12 (0.14)	0.62 (0.19)	1.99 (0.02)	1.22 (0.03)	0.69 (0.17)
K_s (nM)	10.4 (2.9)	4.1 (2.8)	240 (3.6)	46.6 (1.2)	17.2 (3.5)
K_p (nM)	6.4 (2.3)	3.3 (3.3)	79.6 (2.6)	19.6 (0.8)	12.3 (3.1)

D)

SUMO2 Substrate				
	SENP1		SENP2	
[NaCl] mM	20	100	20	100
k_{cat} (s^{-1})	1.62 (0.14)	1.76 (0.85)	1.12 (0.46)	1.16 (0.19)
K_s (nM)	21.0 (2.3)	22.9 (1.1)	3.4 (0.2)	28.0 (0.5)
K_p (nM)	18.5 (6.4)	14.8 (0.7)	4.6 (0.5)	62.9 (3.1)

Figure 2-5 Kinetic characterization summary.
Summary tables of determination of kinetic constants for SENP1 WT, SENP1 mutants and SENP2 WT in different salt conditions and with different SUMO substrates. The standard deviation of the parameters are indicated in parenthesis.

shows a change from nearly 108×10^6 $M^{-1}s^{-1}$ at low salt to 2.7×10^6 $M^{-1}s^{-1}$ at high salt. Elucidation of these kinetic constants further supports the activity-based observations from the ratiometric studies with the two enzymes when tested under a spectrum of salt concentrations.

The dramatic influence of ionic strength supports the view that the electrostatics of the environment are important in SENP function. Considering the salt-based results, the structures of SENP1 WT and SENP2 WT were compared to identify residues near the SUMO binding site that differ between the two in charge. Two lysine residues in SENP1 were identified that are a glycine in SENP2 in comparison. In essentially making single residue chimeras within SENP1, a series of lysine to glycine SENP1 mutants were generated and tested for kinetic characterization. Kinetic analysis was performed on the catalytic domains of SENP1 WT, SENP2 WT, SENP1 K455G, SENP1 K515G, and the double mutant SENP1 K455,515G at their optimized salt conditions (Figure 2-5C). These results show that loss of K455 improves substrate and product binding affinity while slowing k_{cat} in comparison to SENP1 WT. In contrast, the other single lysine mutant, K515G, shows a substantially reduced substrate and product binding affinity and enhanced k_{cat} while the double lysine mutant's kinetic constants are in between these two individual mutants. These data suggest that the loss of lysine at the 455 position improves substrate and product binding, but the higher affinity binding makes it more difficult for enzyme to shed the product at the expense of a loss in k_{cat}. The loss of substrate affinity in the SENP1 K515G mutant to a value beyond the affinity that is found with SENP2 suggest that while this lysine is important for substrate binding in SENP1, there are other residues in SENP2 that overcome the effect of the loss of that lysine to maintain better affinity for substrate.

Additionally, a SUMO2 version of the FRET substrate was also tested at low and high salt for SENP1 WT and SENP2 WT (Figure 2-5D). For SENP1, protease activity against SUMO2 seems to be unaffected by salt with respect to k_{cat}, K_s, and K_p. In comparing SENP1's kinetic constants at 20 mM salt for SUMO1 versus SUMO2, SUMO1 has a lower turnover but enhanced substrate affinity and apparent catalytic efficiency (k_{cat}/K_s at 77×10^6 $M^{-1}s^{-1}$). For SENP2, the k_{cat} is unaffected by salt with SUMO2 processing, but the affinity for substrate and product and apparent catalytic efficiency are reduced at 100 mM salt. Additionally, in comparing SENP2 with SUMO1 versus SUMO2 processing, at high salt the k_{cat} is improved for SUMO2 while the apparent catalytic efficiencies are the same (SUMO1: 40×10^6 $M^{-1}s^{-1}$, SUMO2: 41×10^6 $M^{-1}s^{-1}$). In contrast, at low salt, SENP2's apparent catalytic efficiency for SUMO2 is nearly 60-fold greater than for SUMO1. Previously reported results on SUMO isoform maturation preferences for SENP enzymes show that SENP1 shows slight preference for SUMO1 over SUMO2/3 but robustly processes all SUMO isoforms (Hickey et al., 2012; Kolli et al., 2010; Mendes et al., 2016; Mikolajczyk et al., 2007; Shen et al., 2006). In contrast, while SENP2 can also process all three SUMO isoforms, it shows a substantial preference for SUMO2/3 over SUMO1. These results show that at low salt, the apparent catalytic efficiencies (k_{cat}/K_s) are consistent with the published SENP preferences for SUMO isoform maturation. In contrast, at the higher salt concentrations, which are in closer proximity to physiologically relevant salt concentrations, SENP1 shows an apparent preference for SUMO2 and SENP2 shows no preference for SUMO isoforms.

4. Discussion

Development of a robust assay for assessing SENP protease activity is essential for the future discovery of inhibitors as well as the characterization and guided further derivatization of hits. In this work, a FRET-optimized substrate capable of high-yield purification was generated. This substrate, SFCypet-SUMO1-SFYpet, is a powerful tool for both screening for potential inhibitors and detailed kinetic analysis. In a high throughput screening setting, a ratiometric approach to the analysis allows for the robust yet simple survey of large libraries of compounds. With the small volumes used in 384 and 1536-well screening formats, the ratiometric approach is able to accommodate potential inconsistencies in protein concentration from well to well. This format also allows for higher throughput in testing compounds in a dose-responsive and time course dependent manner to determine properties related to inhibitor-target selectivity and structure activity relationships. Additionally, with the initial fluorescence and through monitoring their signals over time, this approach allows for detection of compounds interfering with substrate fluorescence and potentially giving false positives for inhibition. These attributes along with the tolerance to DMSO make the ratiometric FRET assay a powerful tool for future exploration of SENP1 inhibitors.

While the ratiometric approach has tremendous utility, the ratio is not directly proportional to the reaction rate. The limitations of the ratiometric approach arise from a low channel 1 dynamic range that limits the minimum concentration of substrate that can be assayed, and the compounding fluorescent signals due to continued channel 2 signal from direct SFYpet stimulation in both intact substrate and cleaved product. Using analysis specifically based on channel 2 readout allows for detailed determination of kinetic

constants of SENP enzymes. These details breakdown which particular component(s) are affected under changing enzyme constructs and/or assay conditions. Using this approach, the observations found in the ratiometric approach on the differential effects of salt on enzyme activity can be further understood. Beyond the changes in activity, the effects of salt specifically on substrate binding vs enzyme catalytic activity can be parsed out. The assessment of SENP1 mutants shows the additional utility of the channel 2 approach to probe specific residues for their individual roles in substrate binding and catalytic activity. The further evidence that mature free SUMO1 functions as an inhibitor in the assay supports our hypothesis on the significant product inhibition observed. Previous reports on the development of assays for monitoring SENP activity have failed to account for this product inhibition despite its impact. Additionally, the mature free SUMO1 work shows that we can determine the mechanism of inhibition of future identified inhibitors in an unbiased manner and the residues responsible for inhibitor binding and selectivity can also be potentially probed. As opposed to the steady-state assumptions and use of initial rate estimates to determine kinetic constants, the global fitting of the entire reaction progress curves allows for inclusion of the array of data points that sample near critical points for determining the constants. Within the matrix of conditions, each well's conditions change throughout the progress of the reaction, so fitting the entire curve increases the robustness of extraction of kinetic constants.

Within the properties of the assays, the finding that SENP activity is highest when the pH is more basic (pH 8-9) is supportive of the mechanism of a cysteine protease. In the Asp-His-Cys triad, the aspartate hydrogen-bonds with or extracts a proton from histidine, which then bonds with Cys' thiol hydrogen to generate a stronger nucleophile at

the cysteine for eventual peptide bond hydrolysis. In basic environments that still preserve SENP protein structure, the potential protonation of the nucleophilic thiol by a hydronium ion is limited as compared to acidic environments.

As another item related to the properties and optimization of the assay, the differential sensitivity to salt is a unique finding to the SUMO protease field. Others who have used the same catalytic domains to study SENP1 and SENP2 activity have unknowingly tested SENP2 at its suboptimal conditions to show the selectivity of any substrate or inhibitor for a SENP paralog. The evidence presented here shows that in testing the enzymes under the same conditions, selectivity in substrate or enzyme binding may be a product of a number of factors, including natural differences in ideal microenvironment at the site of catalysis instead of specifics regarding actual substrate or inhibitor binding. This could lead researchers astray in leveraging perceived structure activity relationships to attempt to generate inhibitors more selective for a SENP paralog. With the SFCypet-SUMO1-SFYpet as a substrate and the concurrent optimization of assay conditions and analysis to cater to different needs (structure activity relationships versus k_{cat} for example), we have developed a robust tool amenable to meeting the evolving demands of a small molecule inhibitor campaign.

5. References

1. Chen, C.-H., Namanja, A.T., and Chen, Y. (2014). Conformational flexibility and changes underlying activation of the SUMO-specific protease SENP1 by remote substrate binding. Nat. Commun. *5*.

2. Eifler, K., and Vertegaal, A.C.O. (2015). SUMOylation-Mediated Regulation of Cell Cycle Progression and Cancer. Trends Biochem. Sci. *40*, 779–793.

3. Hickey, C.M., Wilson, N.R., and Hochstrasser, M. (2012). Function and regulation of SUMO proteases. Nat. Rev. Mol. Cell Biol. *13*, 755–766.

4. Jiang, L., Liu, Y., Song, Y., Saavedra, A.N., Pan, S., Xiang, W., and Liao, J. (2013). Internal Calibration Förster Resonance Energy Transfer Assay: A Real-Time Approach for Determining Protease Kinetics. Sensors *13*, 4553–4570.

5. Jiang, L., Saavedra, A.N., Way, G., Alanis, J., Kung, R., Li, J., Xiang, W., and Liao, J. (2014). Specific substrate recognition and thioester intermediate determinations in ubiquitin and SUMO conjugation cascades revealed by a high-sensitive FRET assay. Mol. BioSyst. *10*, 778–786.

6. Kang, X., Qi, Y., Zuo, Y., Wang, Q., Zou, Y., Schwartz, R.J., Cheng, J., and Yeh, E.T.H. (2010). SUMO-Specific Protease 2 Is Essential for Suppression of Polycomb Group Protein-Mediated Gene Silencing during Embryonic Development. Mol. Cell *38*, 191–201.

7. Kolli, N., Mikolajczyk, J., Drag, M., Mukhopadhyay, D., Moffatt, N., Dasso, M., Salvesen, G., and Wilkinson, K.D. (2010). Distribution and paralogue specificity of mammalian deSUMOylating enzymes. Biochem. J. *430*, 335–344.

8. Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-Sequence Determinants of Gene Expression in Escherichia coli. Science *324*, 255–258.

9. Leach, C., Tian, X., Mattern, M., and Nicholson, B. (2009). Detection and Characterization of SUMO Protease Activity Using a Sensitive Enzyme-Based Reporter Assay. In SUMO Protocols, H. Ulrich, ed. (Humana Press), pp. 269–281.

10. Leippe, D.M., Nguyen, D., Zhou, M., Good, T., Kirkland, T.A., Scurria, M., Bernad, L., Ugo, T., Vidugiriene, J., Cali, J.J., et al. (2011). A bioluminescent assay for the sensitive detection of proteases. BioTechniques *51*, 105–110.

11. Liu, Y., and Liao, J. (2013). Quantitative FRET (Förster Resonance Energy Transfer) Analysis for SENP1 Protease Kinetics Determination. J. Vis. Exp. JoVE.

12. Liu, Y., Song, Y., Madahar, V., and Liao, J. (2012). Quantitative Förster resonance energy transfer analysis for kinetic determinations of SUMO-specific protease. Anal. Biochem. *422*, 14–21.

13. Liu, Y., Shen, Y., Zheng, S., and Liao, J. (2015). A novel robust quantitative Förster resonance energy transfer assay for protease SENP2 kinetics determination against its all natural substrates. Mol. Biosyst. *11*, 3407–3414.

14. Madu, I.G., and Chen, Y. (2001). Assays for Investigating deSUMOylation Enzymes. In Current Protocols in Molecular Biology, (John Wiley & Sons, Inc.), p 10.30.1-10.30.13.

15. Martin, S.F., Hattersley, N., Samuel, I.D.W., Hay, R.T., and Tatham, M.H. (2007). A fluorescence-resonance-energy-transfer-based protease activity assay and its use to monitor paralog-specific small ubiquitin-like modifier processing. Anal. Biochem. *363*, 83–90.

16. Martin, S.F., Tatham, M.H., Hay, R.T., and Samuel, I.D.W. (2008). Quantitative analysis of multi-protein interactions using FRET: Application to the SUMO pathway. Protein Sci. *17*, 777–784.

17. Mendes, A.V., Grou, C.P., Azevedo, J.E., and Pinto, M.P. (2016). Evaluation of the activity and substrate specificity of the human SENP family of SUMO proteases. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1863*, 139–147.

18. Mikolajczyk, J., Drag, M., Békés, M., Cao, J.T., Ronai, Z. 'ev, and Salvesen, G.S. (2007). Small Ubiquitin-related Modifier (SUMO)-specific Proteases PROFILING THE SPECIFICITIES AND ACTIVITIES OF HUMAN SENPs. J. Biol. Chem. *282*, 26217–26224.

19. Nacerddine, K., Lehembre, F., Bhaumik, M., Artus, J., Cohen-Tannoudji, M., Babinet, C., Pandolfi, P.P., and Dejean, A. (2005). The SUMO Pathway Is Essential for Nuclear Integrity and Chromosome Segregation in Mice. Dev. Cell *9*, 769–779.

20. Nguyen, A.W., and Daugherty, P.S. (2005). Evolutionary optimization of fluorescent proteins for intracellular FRET. Nat. Biotechnol. *23*, 355–360.

21. Nicholson, B., Leach, C.A., Goldenberg, S.J., Francis, D.M., Kodrasov, M.P., Tian, X., Shanks, J., Sterner, D.E., Bernal, A., Mattern, M.R., et al. (2008). Characterization of ubiquitin and ubiquitin-like-protein isopeptidase activities. Protein Sci. *17*, 1035–1043.

22. Orcutt, S.J., Wu, J., Eddins, M.J., Leach, C.A., and Strickler, J.E. (2012). Bioluminescence Assay Platform for Selective and Sensitive Detection of Ub/Ubl Proteases. Biochim. Biophys. Acta *1823*, 2079–2086.

23. Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T.C., and Waldo, G.S. (2006). Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. *24*, 79–88.

24. Segel, I.H. (1975). Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems (J. Wiley).

25. Sharma, P., Yamada, S., Lualdi, M., Dasso, M., and Kuehn, M.R. (2013). Senp1 Is Essential for Desumoylating Sumo1-Modified Proteins but Dispensable for Sumo2 and Sumo3 Deconjugation in the Mouse Embryo. Cell Rep. *3*, 1640–1650.

26. Shen, L.N., Dong, C., Liu, H., Naismith, J.H., and Hay, R.T. (2006). The structure of SENP1-SUMO-2 complex suggests a structural basis for discrimination between SUMO paralogues during processing. Biochem. J. *397*, 279–288.

27. Wang, L., Wansleeben, C., Zhao, S., Miao, P., Paschen, W., and Yang, W. (2014). SUMO2 is essential while SUMO3 is dispensable for mouse embryonic development. EMBO Rep. *15*, 878–885.

Chapter 3

High Throughput Screening for SENP1 inhibitors and characterization of identified small molecule hits

1. Introduction

Based on the coverage of SUMO-based post-translational modifications (PTMs) in the proteome, SUMO is considered to be involved in most if not all cellular processes. Consistent with this integral role, SUMO has been implicated in numerous diseases, including many cancers, diabetes, neurodegenerative diseases, and vascular disease. Consequently, the SUMO pathway has become a target of study for potential therapeutics.

Though there are notable exceptions, the general consensus of SUMO-driven effects on transcription factors and other transcription-modifying factors is that SUMOylation decreases their transcriptional output (Chymkowitch et al., 2015). In the context of the highly proliferative disease of cancer, the subdued induction of transcription by SUMOylation can be a helpful strategy to modulate production of oncogenes. The development and optimization of protein activators is generally considered more difficult

‡ High throughput screening and rank order inhibition with purinergic receptor inhibitors were performed by Dr. Jorge Iñiguez-Lluhí and Angelica Willis. Carrie M. Johnson performed the dose-response testing of HTS hits, class 8 characterization, orthogonal inhibition with PPNDS and cell extracts, PPNDS characterization, and PPNDS time dependence. In collaboration with Dr. Jorge Iñiguez-Lluhí and Melissa Lemke, Carrie M. Johnson performed Class 1 characterization.

than inhibitors (Darby et al., 2014; Zorn and Wells, 2010). Instead of trying to activate the SUMOylation E1, E2, E3 pathway, efforts targeting the SUMO system have been geared towards generating inhibitors for SENP enzymes, as they are responsible for deSUMOylating modified proteins. As a result of this inhibition, levels of SUMOylation and the resulting subdued transcriptional activity are maintained.

A number of groups have targeted the SENP enzymes for inhibition. Initial work in the study of SENPs used a SUMO construct containing an electrophilic trap of a vinyl sulfone off its C-terminal glycine. This construct formed covalent adducts to the thiol group of the catalytic cysteine in SENP enzymes to inhibit their protease activity (Hemelaar et al., 2004). Using the C-terminal tail residues of SUMO as a guide, others have used short peptidyl chains with vinyl sulfones (Borodovsky et al., 2005). Because the vinyl sulfone is a highly reactive Michael acceptor, these inhibitors are nonspecific and capable of inhibiting the SUMOylation E1 and E2 machinery along with other potential reactive cysteines. Following this lead, a short SENP specific peptide inhibitor equipped with a different electrophilic trap, a fluoromethylketone, was used to inhibit SENP1 and SENP2 (Dobrotă et al., 2012). This functional group is smaller and better mimics the glycine-like size limitations imposed by the tryptophan tunnel covering the catalytic cysteine. Though selective for SENP2 over SENP1, this probe still had high levels of background labeling and formed covalent adducts with SENP6, SENP7, and other species in the cellular milieu.

Aside from peptides, small molecules and peptidomimetic inhibitors have been developed to inhibit SUMO proteases. Starting with the hits from a screen of a library of known cysteine protease inhibitors, analogs have been synthesized to generate inhibitors with micromolar-potency for SENP1 and its closest related homolog, SENP2, with

increasing specificity and improving pharmacokinetic properties. Some of these inhibitors were equipped with reactive aza-epoxide or acyloxymethyketone (AOMK) groups (Albrow et al., 2011). Acting as a Michael acceptor, the AOMK warhead is glycine-like in size and able to enter the SENP tryptophan tunnel. These AOMK-based inhibitors were shown to selectively form covalent adducts with SENP enzymes, even when added to complex protein mixtures. Deviations away from the peptide backbone with enhanced hydrophobicity have been employed to enhance the selectivity and stability of inhibitors. Qiao *et al*. added a benzodiazepine functionality with an aldehyde off the central ring to a short peptidomimetic linker to create a family of compounds with better pharmacokinetic properties. These inhibitors showed low micromolar IC_{50} values against SENP1 in purified protein assays and notably inhibited cancer cell growth *in vitro* at similar micromolar concentrations (Qiao et al., 2011). Using a synthesized library of phenylurea derivatives known to inhibit HIF1α, Uno *et al* achieved selectivity for SENP1. Using an *in vitro* fluorogenic SUMO1-AMC (7-amino-4-methylcoumarin) assay they were able to achieve IC_{50} values as low as 29.6 μM for SENP1 while SENP2 showed no significant inhibition up to 100 μM (Uno et al., 2012).

Taking advantage of growth in computational modeling and dynamics, virtual screening for SENP inhibitors has also been employed. By using known crystal structures of SENP1 in complex with SUMO, labs have been able to efficiently screen hundreds of thousands of compounds and then assay top hits in biological assays to confirm the *in silico* findings. Using the Glide program and then docking the confirmed hits to understand their binding mode near the SENP active site, Chen *et al*. used rational drug design to attempt to further enhance the potency of potential inhibitors. Their inhibitors were tested for

SENP1 inhibition using a gel-based isopeptidase assay with RanGAP-SUMO2 substrates and showed IC_{50} values ranging as low as 1 to 2 μM (Chen et al., 2012). Among other virtual screening efforts targeting SENP1, new chemotypes for inhibitors have arisen, including noncovalent and non-competitive inhibitors (Kumar et al., 2014; Madu et al., 2013; Wen et al., 2014; Zhao et al., 2016). While various assays were used to validate these hits, including NMR, gel-based, and the SUMO-CHOP reporter assay, none showed specificity to SENP1 over SENP2 when tested and only one showed cell permeability.

Recently, Wu *et al*. identified the natural product Momordin Ic as a SENP1 inhibitor. In addition to showing *in vitro* activity against SENP1 in purified proteins and mammalian cancer cell culture, they went beyond previous SENP1 inhibitor exploration to demonstrate inhibitory activity *in vivo*. Using a tumor xenograft model of PC3 cells in Balb/c nude mice, intraperitoneal injection of Momordin Ic slowed tumor growth both by PCNA staining and tumor volume (Wu et al., 2016). The tumor cells treated with Momordin Ic also showed increased apoptosis via terminal deoxynucleotide transferase dUTP nick-end labeling (TUNEL) staining and increased levels of SUMO1 and SUMO2/3 staining indicating accumulating SUMO-conjugates. Using a gel-based assay with RanGAP1-SUMO2 substrate, the IC_{50} was 15.4 μM but mechanism of inhibition was not determined.

Of all the designed SENP1 inhibitors thus far, the rationale behind the drug design was focused on targeting the active site cysteine tunnel. With either peptidomimetics using the SUMO C-terminal tail as inspiration, electrophilic warheads selective for cysteine proteases, or virtual screening focused on the active site, very little if any weight is given to the larger surface of interaction between the globular β-grasp domain of SUMO and the

SENP catalytic domain. This region of substantial protein-protein interaction drives much of the activation of SENP enzymes through long-range allostery and is considered to be another potential area to target in generating SENP inhibitors (Chen et al., 2014; Kumar and Zhang, 2013; Shi et al., 2013).

Using the ratiometric FRET-based assay described in the second chapter, our group engaged in a high throughput screening (HTS) campaign to identify and characterize SENP1 inhibitors. At the outset, it is important to note that the primary screening described below was initiated prior to the start of my dissertation work and multiple individuals contributed to the data in Figure 3-1. This data and corresponding experimental details are included here for the sake of completeness and clarity. Because the full SUMO1 protein is present in the FRET substrate, the assay we developed is capable of detecting potential small molecule inhibitors that function by disruption of SUMO binding at its C-terminal tail or β-grasp domain. The screening assay demonstrated appropriate performance needed for a quality HTS, including tolerance to DMSO, good reproducibility, low cost, and a robust Z' factor (Zhang et al., 1999). Kinetic time course and dose-responsive monitoring allows for detection of non-specific, interfering compounds. Additional measures were taken to prevent redox cycling and generation of reactive oxygen species that could generate false positives by using β-mercaptoethanol instead of dithiothreitol (DTT) and the addition of catalase to the reaction buffer (Johnston, 2011). Following careful confirmation of molecular hits obtained in HTS, commercial derivatives of the parent compound were characterized for their inhibitory properties towards the catalytic domain of SENP1 as well as the full-length enzyme. The mechanism of action of exemplars of active scaffolds were characterized for their mechanism of action and the basis for their selectivity towards

SENP1 catalytic domain (CD) versus SENP2 CD was explored through the analysis of specific mutants.

2. Materials and Methods

High throughput screening. The screening campaign at the University of Michigan Center for Chemical Genomics involved a primary screen endpoint assay in 384 well format (20 μL total volume). Compounds (20-50 μM final concentration) were distributed to assay plates via a pin tool and pre-incubated with SENP1 catalytic domain for 10 minutes. Assays were initiated by addition of SUMO1 FRET substrate and incubated for 30 minutes at room temperature before reading in an Envision plate reader with CFP and YFP filter sets. The positive control for inhibition was absence of enzyme and negative control was DMSO vehicle control. Hits from the primary screen were then re-tested in triplicate in a time course manner to test for spectral interference and reproducibility. Confirmed compounds were then tested in an 8 point (4-150 μM) dose-response in duplicates. Compounds devoid of significant interference and displaying Hill coefficients near unity were ranked by potency. This set was assessed against the CCG database of HTS results from other screens as well as for Pan Assay Inhibitor (PAINS) chemical properties for detection of promiscuous inhibitors. The compounds were then grouped by structural similarity and in consultation with our Medicinal Chemistry collaborator, Dr. Scott Larsen, the most promising classes based on physicochemical properties and synthetic tractability were pursued further. Commercially available compounds and selected analogs were repurchased for end point and dose response testing from fresh powders.

The screening campaign of the NIH molecular Libraries Program collection was carried out at the Sanford-Burnham as part of the NIH MLPCN initiative. The primary screening was done via an endpoint assay in a 1536 well format at a volume of 6.03 μL using the *in vitro* FRET assay buffer conditions described above with 333.3 U/mL catalase added. In each well, 10 μM of respective compound incubated with 20 pM SENP1 CD enzyme and the reactions were initiated by addition of substrate at a final concentration of 25 nM and incubated for 50 min. Plates were read on a PerkinElmer Envision Plate reader equipped with CFP and YFP filters. Positive control for inhibition was absence of enzyme and negative control was compound DMSO vehicle control. Active compounds were re-tested in quadruplicate in a time course and endpoint assay. Confirmed hits from the primary screen were subjected to cheminformatics filtering to remove promiscuous inhibitors in other screens, overtly reactive compounds and those displaying PAINS properties. The resulting set was assayed in a dose-response format in both a time course and endpoint format. Compounds displaying significant potency were then subjected to a confirmatory order of addition test in which compounds were added after completion of the reaction to rule out spurious artifactual effects.

In Vitro **FRET assay for confirmation and characterization**. Reactions were 100 μL final and performed in high or low salt buffer solution (20 mM TrisCl pH 8.0, 250 μM Na·EDTA, 20 mM NaCl or 100 mM NaCl, 10 mM β-Mercaptoethanol, 0.1 mg mL^{-1} bovine serum albumin, 0.01% NP-40). Reactions were initiated by the addition of substrate and fluorescence was read every minute on a Spectramax M5 plate reader excitation at 405 nm and emission measured at 475 and 525 nm.

USP2 CD construction, expression, and purification. USP2 (AA 259-605) was subcloned into pET24a vector (Novagen). Rosetta2 (DE3) cells were grown under kanamycin selection in TB medium at 37 °C until OD_{600} 1.2. Temperature was decreased to 18 °C and isopropyl β-D-1-thiogalactopyranoside was added to 0.2 mM final. Cells were grown for another 20 hours. Harvested cells were resuspended in extraction buffer containing 10 mM Tris pH8.0, 500 mM NaCl, 5% glycerol, 2 mM imidazole, 1 mM BME, 0.5 mM PMSF, 0.5 μL benzonase. Cells were lysed by 3 cycles of French press and clarified by centrifugation at 40,000 × g 40 min at 4°C. 1.5 mL of Ni-NTA agarose were incubated with the clarified supernatant for 1 hour at 4°C rotating. Proteins were eluted by the extraction buffer supplemented with 200 mM imidazole. Pooled fractions were gel-filtered in 50 mM HEPES pH8.0, 150 mM NaCl, 10% glycerol, 0.5 mM Na·EDTA, 1 mM DTT.

PPNDS reversibility. 500 nM SENP1 CD was incubated with 10 μM PPNDS or DMSO vehicle control for 30 minutes at room temperature in reaction buffer (20 mM TrisCl pH 8.0, 250 μM Na·EDTA, 20 mM NaCl, 10 mM β-Mercaptoethanol, 0.1 mg mL^{-1} bovine serum albumin, 0.01% NP-40),. The mixture was then diluted 100-fold and served as the SENP1 enzyme source for the assay, where it made up 10% of the 100 μL final assay volume for an experimental activity concentration of 500 pM. In addition to the enzyme, the final assay solution contained reaction buffer, 600 nM SFCypet-SUMO1-SFYpet substrate, and 10 μM PPNDS or DMSO vehicle control. Reactions were initiated by the

addition of substrate and fluorescence was read every minute on a Spectramax M5 plate reader excitation at 405 nm and emission measured at 475 and 525 nm.

Transfection and preparation of cell extracts. HEK293T cells were grown in DMEM +10% FBS to 70% confluency and seeded in to 10 cm plates at 3×10^6 cells per plate. Cells were incubated overnight. The next morning, the media was exchanged for serum free DMEM and cells were transfected using Lipofectamine 2000 and 5 μg DNA per plate. A pcDNA3 HA SUMO1 plasmid was used for substrate transfection, pCMV6 SENP1 418-C was used for enzyme, and empty pCMV6 for a vector control. After incubating with transfection material for six hours, the media was exchanged for fresh DMEM + 10% FBS and the cells were incubated for 48 hours. Cells were then harvested on ice with lysis buffer (400 mM NaCl, 20 mM HEPES pH 7.4, 5 mM Na·EDTA, 1 mM EGTA, and 1% NP-40). For the vector control and SENP1 extracts the lysis buffer was supplemented with 10 mM DTT. For the extracts serving as substrate source and transfected with HA SUMO1, the lysis buffer was supplemented with 20 mM NEM to alkylate endogenous SUMO proteases and preserve natural SUMOylated substrates. After incubation of the extracts in the NEM supplemented buffer, excess DTT was added to quench any remaining NEM.

Cell extract assay. Total protein concentration of each extract type was determined by Bradford assay. Substrate extracts were mixed with equal total protein content from SENP1 extracts or vector extracts and incubated at room temperature for 15 minutes in a Tris reaction buffer (10 mM TrisCl pH 8.0, 500 μM Na·EDTA, 10 mM β-Mercaptoethanol, 5mM DTT). In the presence of inhibitor, these mixtures were supplemented with 50 μM

PPNDS or equal volume DMSO as a vehicle control. Reactions were quenched with equal volumes of hot 4X SDS sample buffer. Samples were loaded and separated in a 10% SDS-PAGE gel and either stained for analysis or transferred to a nitrocellulose membrane for immunoblot analysis.

Western Blot. Membrane was blocked overnight at 4 °C in 5% nonfat dry milk in Tris-buffered saline with Tween (TBST: 10 mM Tris pH 7.5, 150 mM NaCl, 0.1% Tween-20). The membrane was then rinsed with TBST and a 1:500 solution of mouse anti-HA11 ascites antibody in TBST was applied to the membrane for one hour at 4 °C. The membrane was again rinsed in TBST and a 1:500 solution of goat anti-mouse horseradish peroxidase-conjugated antibody was applied to the membrane for one hour at 4 °C. The membrane was rinsed and imaged using a Li-Cor Odyssey Fc imaging system and enhanced chemiluminescent substrate (Thermo Fisher Scientific SuperSignal West Femto Maximum Sensitivity Substrate). Image was quantified using Image Studio software.

Time dependence and Schiff base reduction by $NaBH_4$. In a 10 μL volume, 150 nM SENP1 418-C was incubated at room temperature with 40 μM PPNDS for 1 hour in a HEPES incubation buffer (20 mM HEPES pH 8.0, 20 mM NaCl, 500 μM Na·EDTA, 0.1 mM TCEP and 0.01% NP-40). Following this incubation, 5 μL of fresh 1 M $NaBH_4$ was added and allowed to incubate for 1 hour. This mixture was diluted 100-fold and functioned as the enzyme source and was diluted 10-fold in to the assay final. Specific activity of each sample was assessed via serial dilutions of the mixture with Cypet-SUMO1-YPet substrate in the FRET assay described above.

3. Results

HTS and identification of SENP1 inhibitors

As indicated above, the screening data in this sub-section is the result of a large effort and is included here for clarity and completeness. Using chemical libraries available at the Center for Chemical Genomics (CCG) at the University of Michigan and at the National Institutes of Health Molecular Libraries Probe Production Centers Network, over 500,000 compounds were screened in total. The overall progressive reduction in candidate compounds after each experimental step of the two campaigns are summarized in Figure 3-1. The primary screening was done in an end point assay to identify inhibitors of SENP1 endopeptidase activity against the SFCypet-SUMO1-SFYpet FRET-based construct described in Chapter 2. Average plate Z' scores for the HTS assay were routinely 0.9 or greater, signifying a robust screening assay with a broad detection range between positive and negative controls (Zhang et al., 1999). With the advantage of such a robust assay and confidence in its performance, a more generous cutoff of compounds showing inhibitory activity were categorized as hits in the primary screen for further analysis. Notably, although the MLPCN primary screening campaign yielded promising candidates, none were proven to be true positives. Active compounds were then subjected to a confirmatory order of addition test in which compounds were added after completion of the reaction (All compounds should be negative). Unfortunately, none of the compounds passed this test since all compounds displayed apparent activity, suggesting that they represented false positives and were not pursued further. The reason for the compounds passing previous tests of interference is unclear. The lower substrate concentration and smaller assay volume utilized in the 1536 well format may have been contributing factors. Of note, none of the

active compounds identified in the CCG collection (see below) were present in the MLPCN NIH collection.

The primary screening of 156,667 compounds at the CCG yielded 2,671 hit compounds, which were further analyzed in triplicate for inhibitory activity in a time course manner to eliminate compounds that cause aggregation, interfere with the fluorescence readout, or possess other non-specific inhibitory activity (Figure 3-1A).

A)

Center for Chemical Genomics screen

Description	Number of compounds tested	Hits
Primary screen: Endpoint	165,667	2,671
Time course in triplicate	2,671	240
Kinetic Dose-response	240	115
Fresh compound time course	115	70
Fresh compound dose-response	70	34

B)

National Institutes of Health MLPCN screen

Description	Number of compounds tested	Hits
Primary screen: Endpoint	365,168	774
Endpoint in quadruplicate/ time course in quadruplicate	774	176/133
Endpoint dose-response/ kinetic dose-response	177	140/177
Change of order of addition	177	0

Figure 3-1 Summary of high throughput screening results.
The narrowing of hit compounds based on confirmation studies described in each table for the high-throughput screens performed at the Center for Chemical Genomics (A) and the National Institutes of Health Molecular Libraries Probe Production Center Network (B).

Confirmed compounds were then tested in a dose-responsive manner to assess for attractive potency and Hill coefficients. This narrowed the candidate compounds to 117. These compounds were then assessed against the CCG database of HTS results and pan-assay interference (PAINS) compounds to discard promiscuous inhibitors (Baell and Holloway, 2010). Hits were also tested for specificity to only the SUMO system by testing against

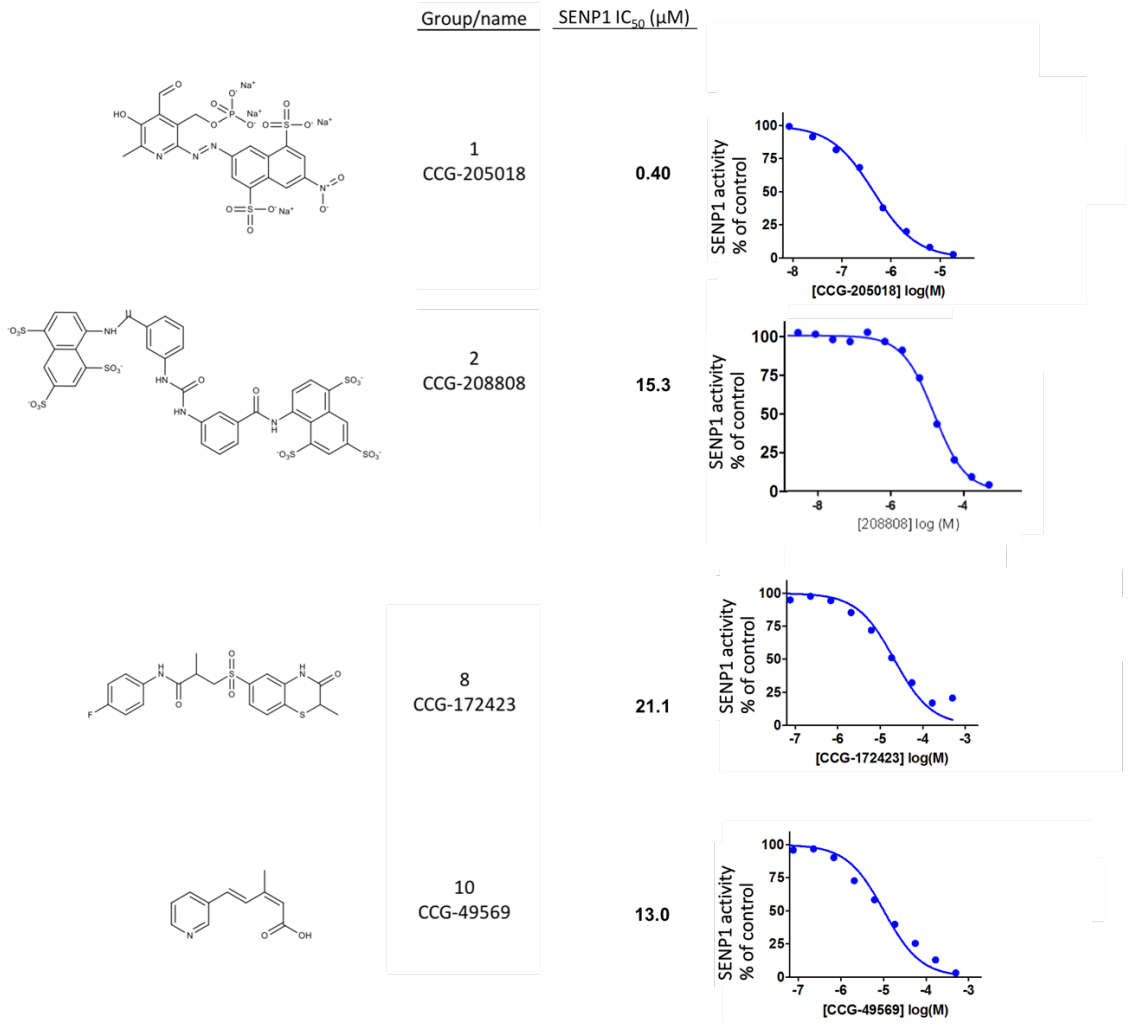


Figure 3-2 A sampling of HTS hits against SENP1.
Structural depiction and dose-response curves of a representative molecule from four classes of compounds showing inhibitory activity against SENP1. Total enzyme activity in the presence of inhibitor is plotted as a percentage of activity normalized to the activity of SENP1 in the presence of DMSO vehicle control.

the ubiquitin protease USP2 and a ubiquitin FRET substrate with the same fluorophores as the SUMO substrate. Commercially available compounds of remaining hits were repurchased for both end point and dose response testing of fresh compound. The list of compound hits was narrowed to 34 and they were structurally grouped into 12 classes of compounds. These remaining compounds were assessed for synthetic chemical tractability and availability of commercial analogs. Four of the classes that were identified for their potency, commercially available analogs, and/or synthetic tractability are depicted in Figure 3-2. PPNDS, a pyridoxal phosphate derivative (Class 1), was the most potent with sub-micromolar potency against SENP1 activity in the FRET assay. The hit in Class 2 corresponds to Suramin, a member of the Azo dye class of compounds. Class 8 and class 10 compounds were identified as classes to pursue initially for further analysis due to their more drug-like characteristics compared to the highly charged class 1 and class 2 compounds.

Characterization of Group 8 and 10 Compounds

Class 8 compounds were determined to be synthetically tractable for future synthetic structural derivatization and had numerous commercially available analogs. Testing a significant number of these commercial analogs suggested a series of structure activity relationships regarding ring sizes and electrochemistry across the pharmacophore scaffold (Figure 3-3A). Many of the substitutions that made the molecule more hydrophobic also enhanced the measured potency of inhibition against SENP1, including increasing size of the A ring and introducing methyl substitutions off the linker region and C ring.

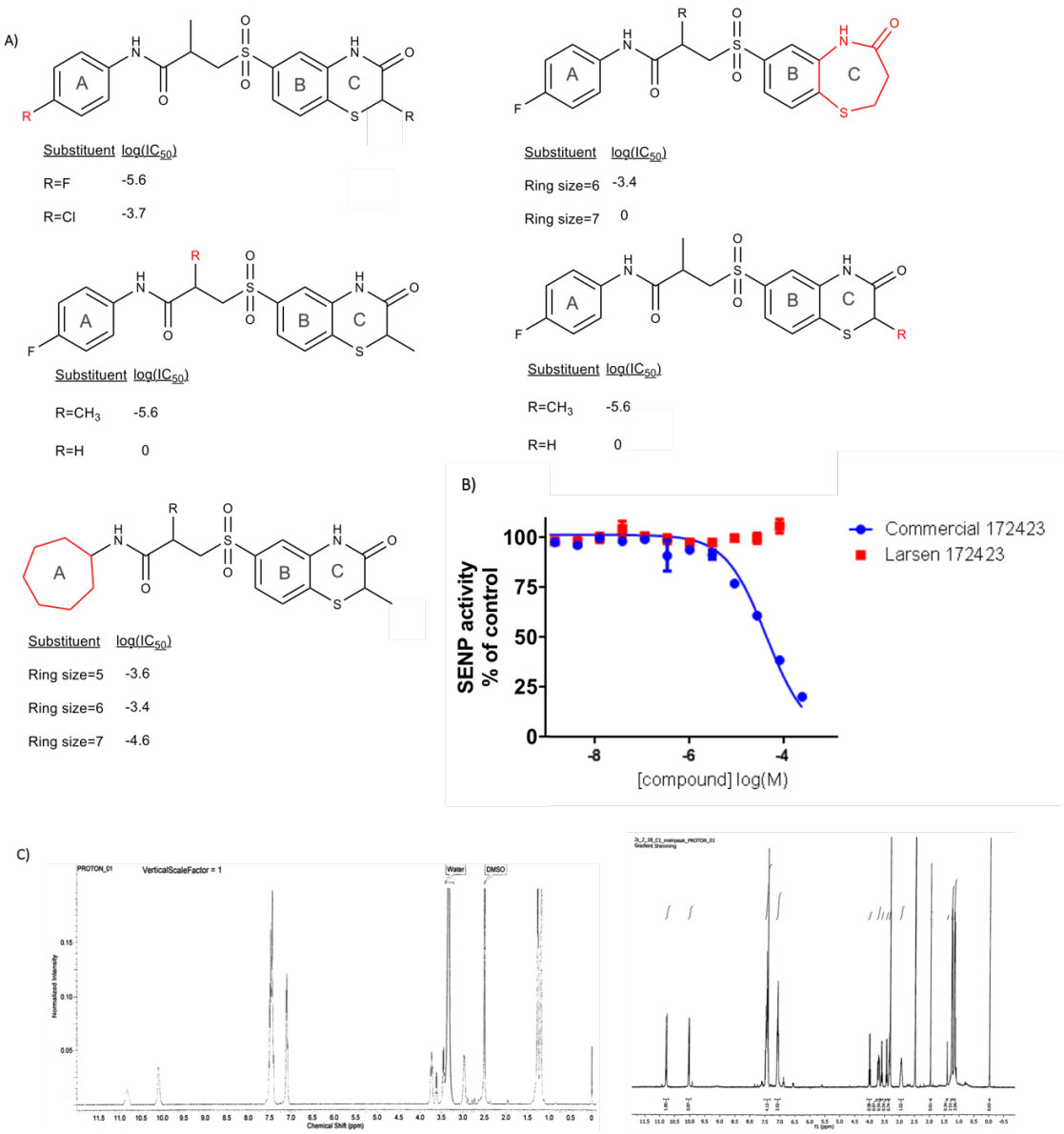


Figure 3-3 Class 8 Characterization.
A) Structure Activity Relationships (SAR) of class 8 compounds. The atom or region colored in red is the part that is changed as described in the quantitative information below each structure. Rings are labeled to ease description of structure activity relationships and activity is presented as the log IC_{50} (M) values of the compound from a dose-responsive assay using the FRET SFCypet-SUMO1-SFYpet substrate. B) Dose-response curve for the purchased commercial CCG-172423 compound (blue) and the same compound synthesized in the lab of Dr. Scott Larsen (red). Total enzyme activity in the presence of compound is plotted as a percentage of activity normalized to the activity of SENP1 in the presence of DMSO vehicle control. C) H^1 NMR spectra of CCG-172423. The spectrum on the left is of the synthesized version from Dr. Scott Larsen's lab and on the right is the purchased commercial version (Performed by Jeffrey Zwicker).

In an effort to leverage the SAR obtained on class 8 compounds, Jeffrey Zwicker, a medicinal chemist in the group of our collaborator, Dr. Scott Larsen, independently developed a synthetic route for synthesizing the most promising candidate from the series (CCG-172423). Unfortunately, this preparation was devoid of any inhibitory activity against SENP1 (Figure 3-3B). Further chemical analysis was performed and carbon and proton NMR spectra for the commercial and independently synthesized preparation yielded identical results (Figure 3-3C). Notably, composition analysis of the commercial material indicated that despite identical organic spectra to the resynthesized sample, the percentage of carbon, oxygen, and nitrogen by weight were below the predicted values (data not shown). This suggests that inorganic impurities present in the commercial preparation may be responsible for the inhibitory activity detected. Unfortunately, a similar effort with class 10 compounds yielded the same negative results.

Class 1 and 2 characterization

Following the failure to confirm the activity of the class 8 and 10 compounds, attention was turned to classes 1 and 2. Notably, both PPNDS (Class1) and Suramin (Class 2) are known inhibitors of purinergic receptors, specifically P2X1 (Coddou et al., 2011a). Using that knowledge, additional known Suramin-like P2X1 receptor inhibitors were assessed for inhibitory activity against SENP1 (Figure 3-4). Despite a lack of functional or structural similarity between SENP1 and P2X1, these class 2 compounds displayed similar rank orders of potency (based on P2X1 published data) albeit with lower overall potency towards SENP1 than P2X1.

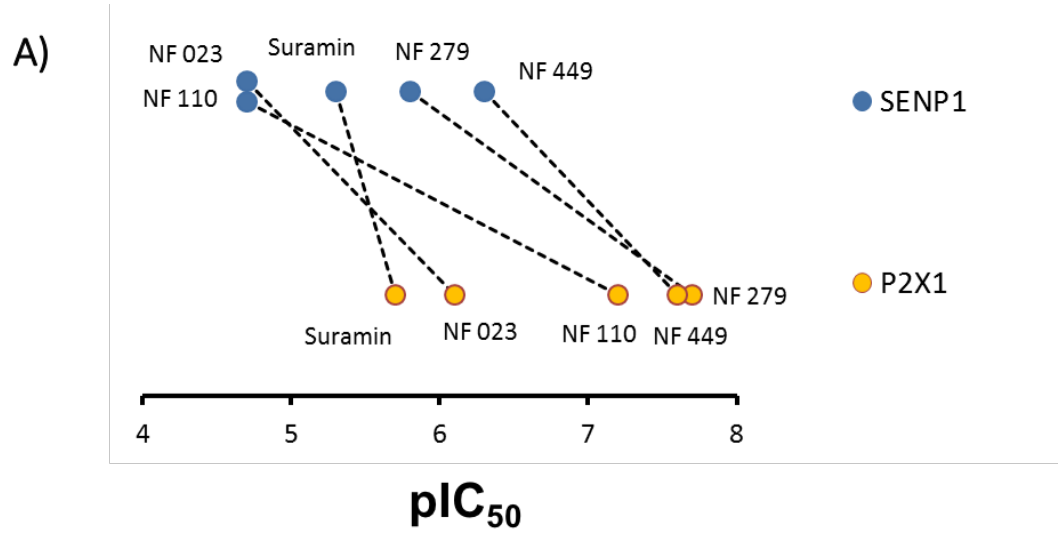


Figure 3-4 Rank order inhibition.
Plot of the -log IC_{50} (pIC_{50}) (M) values for the indicated P2X1 receptor inhibitors in class 2 towards P2X1 (derived from literature) and towards SENP1 in the ratiometric FRET-based assay.

Further characterization of the original class 1 hit, PPNDS revealed a notable selectivity for SENP1 over SENP2 (approximately 20-fold greater potency) when SENP2 was assayed at its optimal salt concentration of 100 mM (Figure 3-5B). When PPNDS inhibition of SENP2 was assayed in the same conditions as SENP1 (20 mM NaCl), no inhibition was observed. Considering the selectivity and potency of PPNDS as a compound from the primary screening library, additional commercially available PPNDS-like compounds were assayed for potential improved potency and potential structure activity relationships (Figure 3-5A). All four compounds contain the pyridoxal phosphate moiety

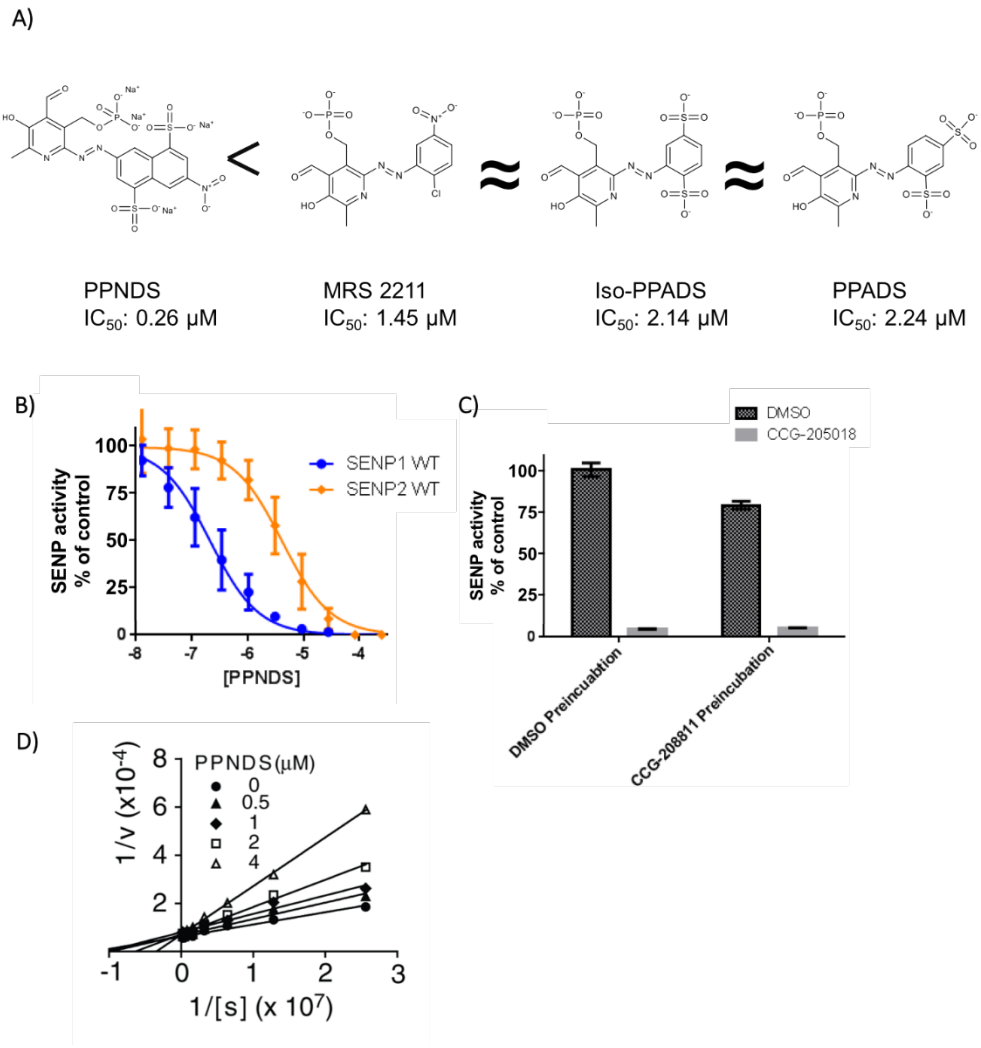


Figure 3-5 Class 1 characterization.
A) Structure activity relationships of commercially available PPNDS-like compounds. IC_{50} values determined via a dose-response assay using the SFCypet-SUMO1-SFYpet FRET substrate and SENP1 catalytic domain (CD). B) Dose-response curve of PPNDS inhibition of SENP1 CD activity vs SENP2 CD activity in the FRET assay. C) To test for reversibility, a plot of SENP1 CD activity after pre-incubation with PPNDS or DMSO vehicle followed by dilution and re-assay in the presence of PPNDS or DMSO. D) Lineweaver-Burke plot of the inverse of substrate concentration vs the inverse of SENP1 CD velocity at different concentrations of PPNDS.

linked to different aromatics appended with charged sulfate groups. PPNDS remained the most potent compound but all four were still more potent than the compounds in the other

classes of identified hits. Although the significant absorbance of PPNDS precluded a full kinetic characterization, reversibility tests and Lineweaver-Burke analysis of initial rates over a modest substrate concentration range suggests that PPNDS acts as a reversible competitive inhibitor (Figure 3-5C and 5D).

To confirm that the inhibitory activity of PPNDS is not a peculiarity of the FRET based assay, an orthogonal assay using cell extracts was employed. As opposed to testing for endopeptidase activity of SENP1 via the FRET-SUMO1 substrate, isopeptidase activity against native substrates could be tested with the cell extracts. Using extracts from cells transfected to overexpress SENP1 or its empty vector as the enzyme source, inhibition of activity by PPNDS (CCG-205018) was tested against extracts of cells transfected with hemagglutinin-tagged SUMO1. Over the course of 15 minutes, the changes in signal from high molecular weight SUMO-conjugated substrates can be observed. In the presence of SENP1 extracts, a substantial loss in high molecular weight HA-SUMO adducts as a percentage of total SUMO signal is seen (Figure 3-6A-B: lane 1, 93% vs lane 2, 12%). Compared to loss in signal in these lanes and the lack of loss in the lanes where vector extract was added, PPNDS inhibited SUMO deconjugation of the natural substrates, as HMW SUMO-conjugates still make up 77% of the total SUMO signal (Figure 3-6A and B: lane 4).

Based on the observed selectivity and mechanism of inhibition evidence, a structural comparison between SENP1 and SENP2 was performed using the published crystal structures of each in complex with SUMO1. Because our work suggests PPNDS was a competitive inhibitor, residues near the SUMO binding site were analyzed. Deviations in protein residue sterics or electronics between the SENP1 and SENP2

catalytic domains were selected as potential target residues for mutagenesis to probe for those responsible for the selectivity in inhibition. Two lysine resides found near the SUMO

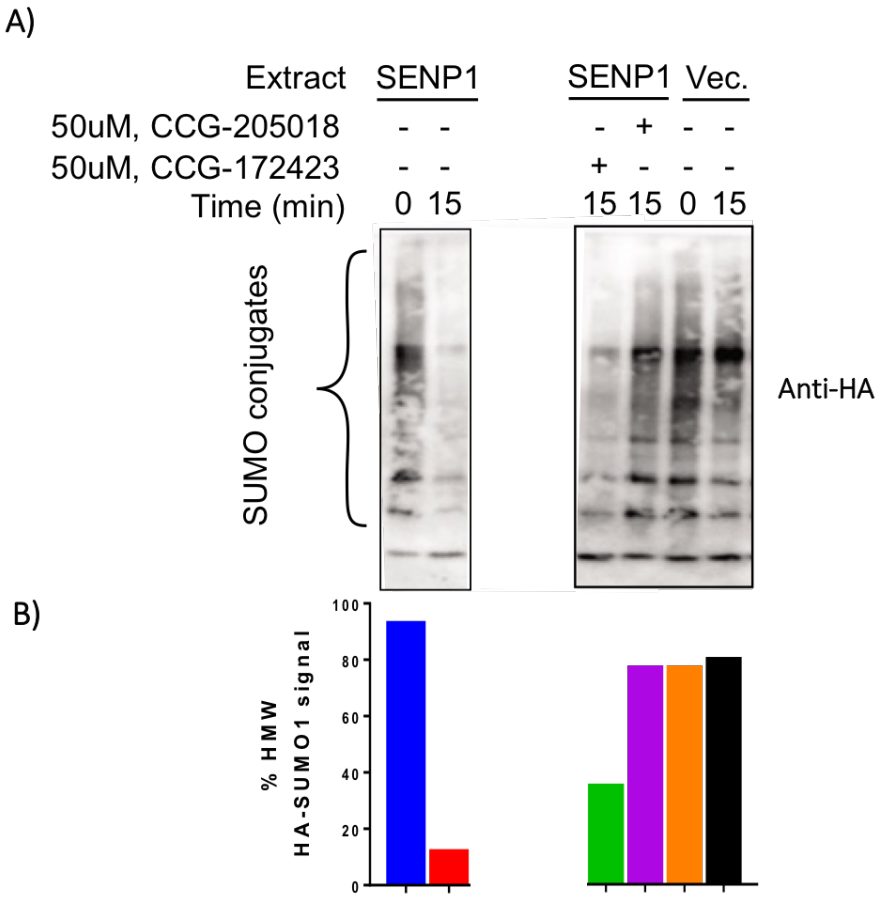
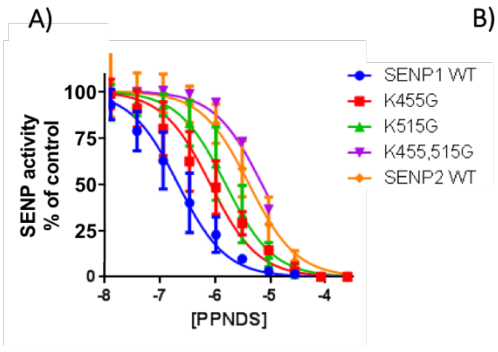
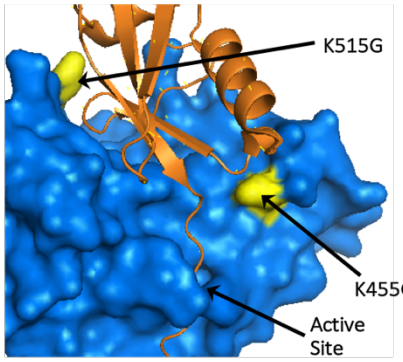


Figure 3-6 Class 1 characterization.
Western blot of cell extracts blotted for hemagglutinin (HA). Every lane contained HA-SUMO1 cell extracts to serve as substrate source for cell extracts transfected to overexpress full length SENP1 or its empty plasmid vector. In addition, each lane also contained 50 µM compound (CCG-205018=PPNDS) or DMSO vehicle control. CCG-1724253 is an inactive compound used as a control. Samples were incubated for 15 minutes after addition of enzyme to substrate extract, and halted via addition of hot SDS sample buffer and boiling of sample. B) Quantification of ratio of HA-signal at high molecular weight (HMW) to total HA signal. The HMW HA signifies SUMOylated substrates, while the total HA signal would include the low molecular weight free SUMO that is released after SENP1 activity. Quantification plot is superimposed below each respective lane described in the western blot.

binding site in SENP1 are absent in SENP2. These two SENP1 residues, K455 and K515 were mutated to the SENP2 equivalent glycine residues both in isolation and in





C)

SENP type	SENP1 (20 mM NaCl) WT	K455G	K515G	K455,515G	SENP2 (100 mM NaCl) WT
k_{cat} (s^{-1})	1.12 (0.14)	0.62 (0.19)	1.99 (0.02)	1.22 (0.03)	0.69 (0.17)
K_s (nM)	10.4 (2.9)	4.1 (2.8)	240 (3.6)	46.6 (1.2)	17.2 (3.5)
K_p (nM)	6.4 (2.3)	3.3 (3.3)	79.6 (2.6)	19.6 (0.8)	12.3 (3.1)
PPNDS IC_{50} (μM)	0.21	0.81	1.50	7.00	4.10
Cheng-Prusoff K_i (mM)	0.020	0.032	1.06	2.23	0.602

D)

Cheng-Prusoff Equation:

$$Ki = \frac{IC_{50}}{1 + \frac{[S]}{Ks}}$$

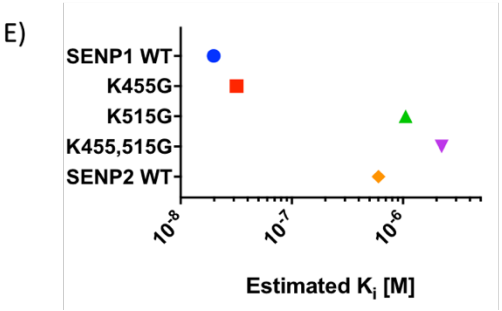


Figure 3-7 Characterization of PPNDS mediated inhibition.
A) Representative dose-response curve of PPNDS-mediated inhibition of activity of SENP1 CD and SENP1 mutants, and SENP2 CD. Activity was normalized to SENP1 CD activity in the presence of DMSO vehicle control. B) Crystal structure complex (2IY1) of SENP1 CD (blue) and SUMO1 (orange ribbon) with the location of SENP1 mutants indicated in yellow and the general region of the active cysteine pointed out for reference. C) Table of kinetic parameters of SENP1 CD, SENP1 mutants, and SENP2 CD. Parameters k_{cat}, K_s and K_p were obtained by global fitting. The IC_{50} values were determined from dose response assays shown in A. Using the Cheng-Prusoff equation (D) and the IC_{50} values for the respective SENP, the inhibitory constant, K_i, was estimated. E) Plot of the estimated Ki values for each SENP as determined by the Cheng-Prusoff equation.

combination (Figure 3-7B). Each single lysine mutant had an IC_{50} value intermediate to the PPNDS IC_{50} value of WT SENP1 and WT SENP2 whereas the double lysine mutant, SENP1 K455,515G, had an IC_{50} value that paralleled that of WT SENP2 (Figure 3-7A). In essence, it appeared that as SENP1 became more SENP2-like with just mutation of the lysine residues near the SUMO binding site, the potency of PPNDS inhibition became on par with SENP2. These results would suggest that the two lysine residues, K455 and K515, are key factors in PPNDS-mediated inhibition. Using the data from the detailed kinetic analysis of the different enzymes (Figure 3-7C), the Cheng-Prusoff equation (Figure 3-7D) was employed to estimate the inhibitory constant. The resulting estimated inhibitory constant, Ki, for each enzyme is found in the table in Figure 3-7C and plotted in Figure 3-7E. These results show that while K455 may contribute to PPNDS-mediated inhibition, K515 seems to play a more significant role, with a binding constant for PPNDS slightly weaker than SENP2 WT.

Throughout PPNDS-based experiments, the potency of inhibitory activity appeared to drift despite using the same stocks of assay components. This generated an investigation into a possible time-dependence of PPNDS-mediated inhibition with respect to the length of time PPNDS was incubated with the enzyme before addition of substrate to begin kinetic assay measurements. In carefully monitoring the time between the initial mixing of PPNDS and SENP1 and the addition of the FRET-SUMO1 substrate, a time-dependent increase in potency was observed (Figure 3-8A). A similar time-dependence was observed in the mutants and SENP2, albeit at different magnitudes of effect (Figure 3-8B). As a result of this observation of time-dependence and the nature of the chemistry of pyridoxal phosphate moieties, the possibility of the formation of a Schiff base between SENP surface lysines

residues was hypothesized as a potential explanation for the effect. The electrophilic carbonyl carbon of the aldehyde group within pyridoxal phosphate is attacked by the free

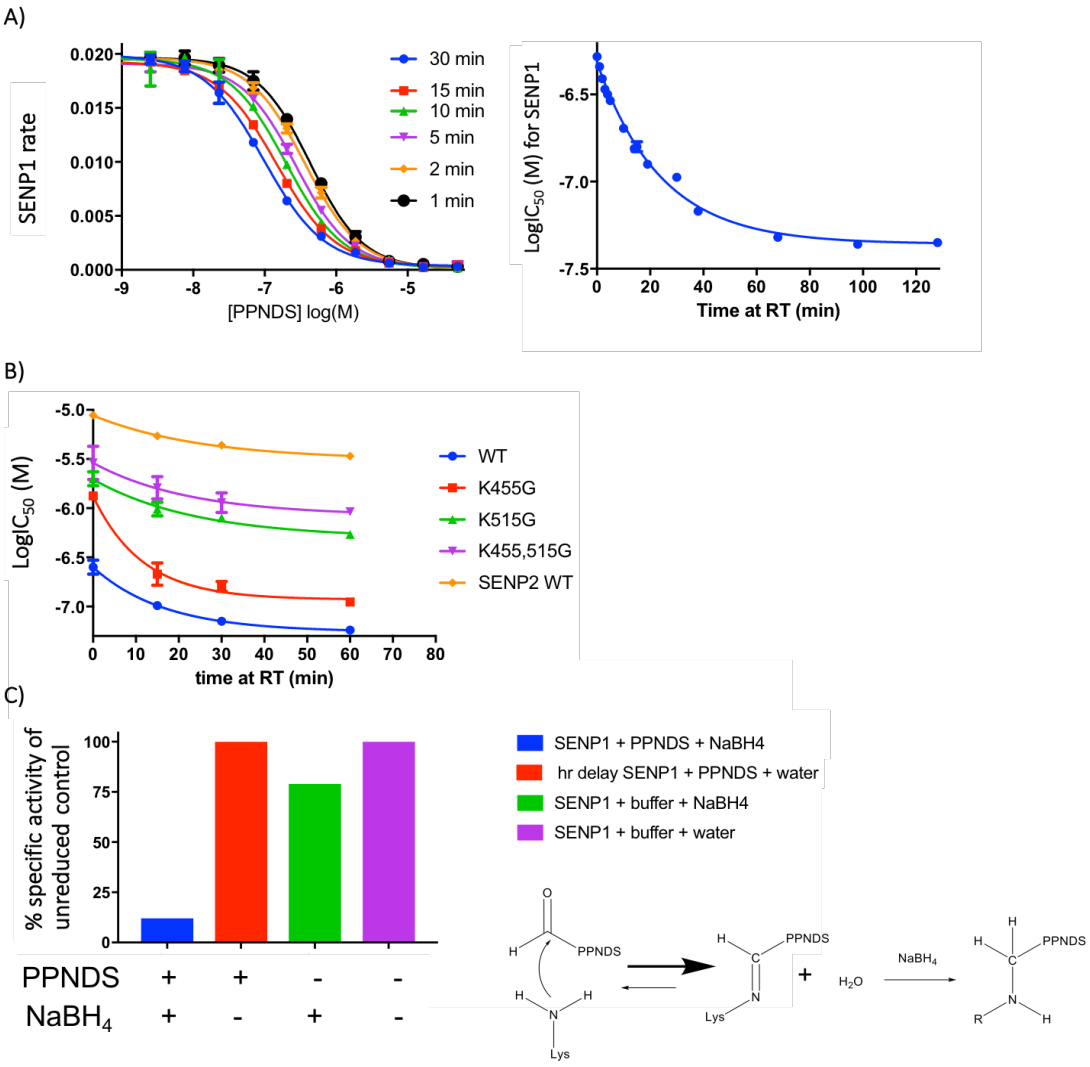


Figure 3-8 PPNDS time dependence.
A) Representative dose-response curves of PPNDS inhibition of SENP1 CD at different times of pre-incubation of SENP1 and PPNDS before addition of substrate to begin the reaction. To the right is a plot of the cumulative and averaged log IC_{50} values plotted with respect to the time of pre-incubation. B) Plot of pre-incubation time versus IC_{50} values for different SENP1 CD mutants and SENP2 CD. C) Following incubation with PPNDS or DMSO vehicle control for 1 hour, samples were incubated in NaBH4 or water vehicle for one hour and diluted 100-fold and subsequently assayed for specific activity. Data was normalized to the activity of the corresponding unreduced samples. To the right is a depiction of the hypothesized chemistry taking place with Schiff-base formation between the PPNDS aldehyde and SENP1 CD lysines with reduction of the reversible reaction via $NaBH_4$.

lone pair of electrons in the protein lysine residue, releasing water and forming a reversible covalent adduct to the enzyme. This bond can be reduced by sodium borohydride to lose the reversibility of the adduct formation and generate a permanent covalent PPNDS-SENP

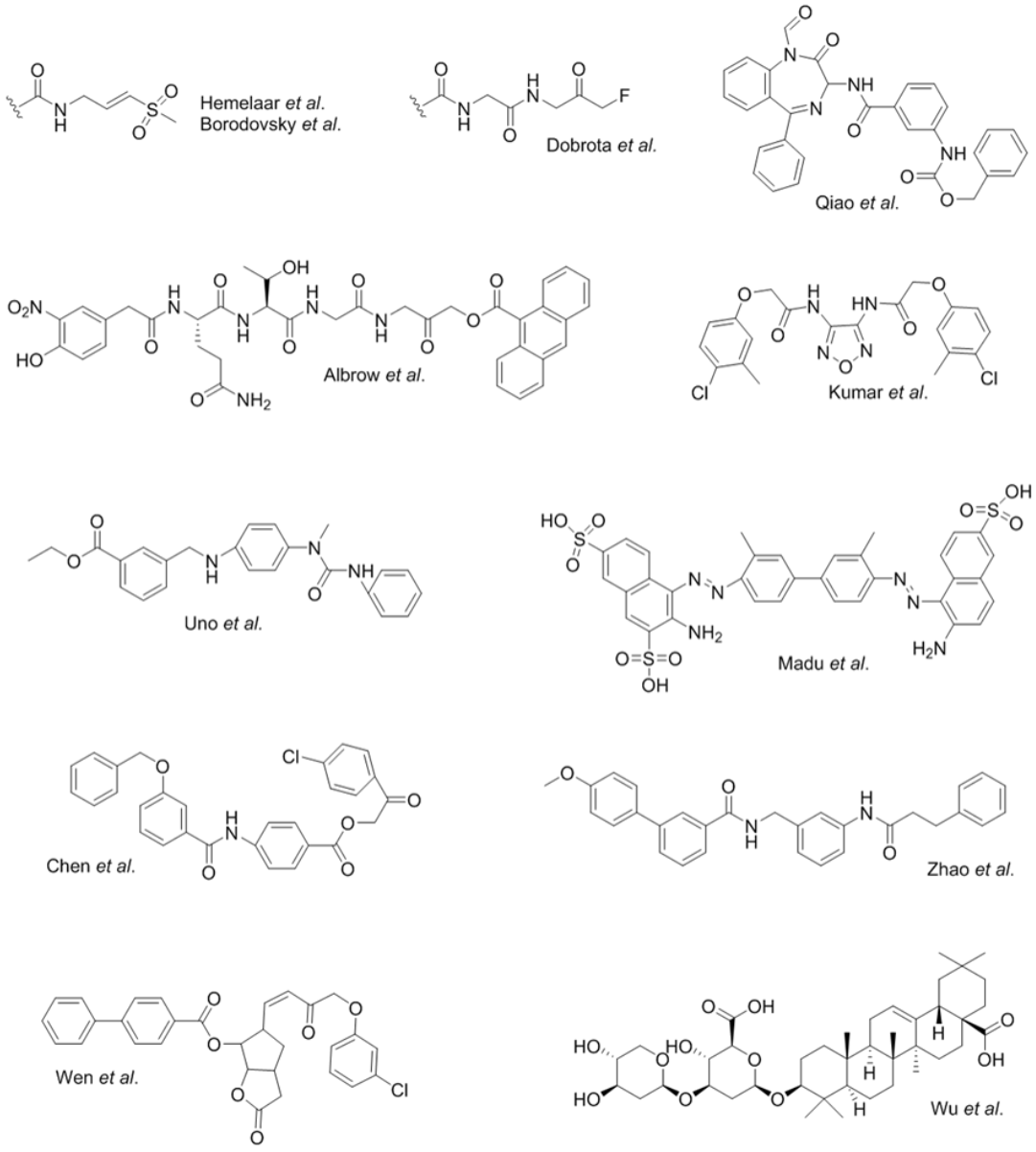


Figure 3-9 Representatives of published small molecule SENP inhibitors.
The published structures of identified SENP inhibitors with the corresponding reference. When applicable, the most potent analog of the class of compounds tested is shown.

adduct. While the data is not definitive, the SENP1 specific activity following incubation with PPNDS and reduction with $NaBH_4$ is severely reduced compared to the untreated control (Figure 3-8C column 1 vs column 3).

4. Discussion

Within the existing published inhibitors for SENP1 and other SENP enzymes, (representatives shown in Figure 3-9), limitations exist with regard to their lack of selectivity, unknown or non-ideal mechanisms of inhibition, and/or lack of robust validation of inhibition of SENPs through orthogonal experiments. Despite the difficulties in generating potent, selective, reliable inhibitors for SENP1, it remains an attractive therapeutic target due its role in numerous disease. In this work, high throughput screening (HTS) with a robust FRET-based assay was employed to discover inhibitors of SENP1 activity. Through numerous steps of refining and triaging the list of potential inhibitors via end-point, kinetic, and dose-response assays, a few classes of compounds were identified for further study. In characterizing compounds for their mechanism of inhibition, their selectivity, and a structural understanding of SENP1-inhibitor interactions, important residues for inhibitor binding and unique pharmacological parallels to unrelated proteins were identified.

The robustness of the ratiometric FRET-based assay was a powerful tool to screen through hundreds of thousands of compounds for not just the initial primary screen, but also in the follow-up work to validate lead compounds and characterize their behavior. Using the assay in a dose response set-up allowed for fast assessment of both the selectivity of compounds for inhibition of SENP1 versus SENP2 and structure activity relationships

among derivatives within compound classes. The ratiometric FRET-based assay allowed for characterization of PPNDS as a competitive, reversible inhibitor that depends on lysine 515 for selective inhibition of SENP1. Importantly, the reliability of the robust assay allowed for observation of the time-dependent changes in PPNDS potency. Had the assay not been reproducible, the drifting in PPNDS potency may have been attributed to alternative confounding variables and not have been noticed. PPNDS chemistry and the time dependence further supported generation of the series of SENP1 lysine mutants to test for potential Schiff base formation. Though not fully conclusive, the data support the Schiff base hypothesis, with some recovery of SENP1 activity observed without addition of the $NaBH_4$ reducing agent. The magnitude of change in potency in the time dependence plots in IC_{50} values for different SENP1 mutants (Figure 3-5B) further supports the Cheng-Prusoff estimated inhibitory constants. The magnitude of the potency change is reduced in SENP1 K515G compared to SENP1 K455G, which might be expected if K515 is the target lysine important for Schiff base formation. Due to the lack of full recovery of activity in non-reduced but PPNDS incubated SENP1 and the continued observance of time dependence, it is likely that PPNDS binds tightly regardless of Schiff base formation but the formation further enhances the potency. Additionally, it is possible that additional or alternative chemistry is taking place that is responsible for the time-dependent inhibition.

Our orthogonal tests with PPNDS using cells extracts and a non-FRET-based approach support our HTS efforts and validation of PPNDS as a SENP1 inhibitor. As an additional advantage, this assay also tests for isopeptidase activity instead of the endopeptidase activity assessed in the single polypeptide SUMO1 substrate used in the FRET-based assays. While the chemistry is similar in both activities in that hydrolysis of

a peptide bond is occurring, the difference in rotational degrees of freedom between the normal polypeptide backbone of SUMO precursor versus the unrestricted rotation in the long side chain of lysines create different potential steric-driven limitations on SENP activity (Komander et al., 2009; Shen et al., 2006). In testing PPNDS in cell extracts, inhibition of deSUMOylation of natural substrates by full length SENP1 shows that PPNDS inhibition is not unique to our synthetic substrate or just the catalytic domain of SENP1. PPNDS can inhibit SENP1 activity against a variety of natural cellular targets.

Both PPNDS (Class 2) and Suramin (Class 1) are published purinergic receptor antagonists, specifically the ionotropic P2X family of receptors (Coddou et al., 2011a, 2011b). A series of other known P2X receptor antagonists were tested against SENP1 and remarkably showed inhibition with some parallel rank order pharmacology in IC_{50} values. Despite no structural or functional similarity between SENPs and P2X receptors, the novel discovery of inhibition for SENP1 by P2X antagonists generates considerable anticipation as to other small molecules to test and suggests a potential mechanism of regulation of SENP1 that warrants future study.

5. References

1. Albrow, V.E., Ponder, E.L., Fasci, D., Békés, M., Deu, E., Salvesen, G.S., and Bogyo, M. (2011). Development of Small Molecule Inhibitors and Probes of Human SUMO Deconjugating Proteases. Chem. Biol. *18*, 722–732.

2. Baell, J.B., and Holloway, G.A. (2010). New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. J. Med. Chem. *53*, 2719–2740.

3. Borodovsky, A., Ovaa, H., Meester, W.J.N., Venanzi, E.S., Bogyo, M.S., Hekking, B.G., Ploegh, H.L., Kessler, B.M., and Overkleeft, H.S. (2005). Small-Molecule Inhibitors and Probes for Ubiquitin- and Ubiquitin-Like-Specific Proteases. ChemBioChem *6*, 287–291.

4. Chen, C.-H., Namanja, A.T., and Chen, Y. (2014). Conformational flexibility and changes underlying activation of the SUMO-specific protease SENP1 by remote substrate binding. Nat. Commun. *5*.

5. Chen, Y., Wen, D., Huang, Z., Huang, M., Luo, Y., Liu, B., Lu, H., Wu, Y., Peng, Y., and Zhang, J. (2012). 2-(4-Chlorophenyl)-2-oxoethyl 4-benzamidobenzoate derivatives, a novel class of SENP1 inhibitors: Virtual screening, synthesis and biological evaluation. Bioorg. Med. Chem. Lett. *22*, 6867–6870.

6. Chymkowitch, P., Nguéa P, A., and Enserink, J.M. (2015). SUMO-regulated transcription: Challenging the dogma. BioEssays *37*, 1095–1105.

7. Coddou, C., Yan, Z., Obsil, T., Huidobro-Toro, J.P., and Stojilkovic, S.S. (2011a). Activation and Regulation of Purinergic P2X Receptor Channels. Pharmacol. Rev. *63*, 641–683.

8. Coddou, C., Stojilkovic, S.S., and Huidobro-Toro, J.P. (2011b). Allosteric modulation of ATP-gated P2X receptor channels. Rev. Neurosci. *22*, 335–354.

9. Darby, J.F., Landström, J., Roth, C., He, Y., Davies, G.J., and Hubbard, R.E. (2014). Discovery of Selective Small-Molecule Activators of a Bacterial Glycoside Hydrolase. Angew. Chem. Int. Ed. *53*, 13419–13423.

10. Dobrotă, C., Fasci, D., Hădade, N.D., Roiban, G.-D., Pop, C., Meier, V.M., Dumitru, I., Matache, M., Salvesen, G.S., and Funeriu, D.P. (2012). Glycine Fluoromethylketones as SENP-Specific Activity Based Probes. ChemBioChem *13*, 80–84.

11. Hemelaar, J., Borodovsky, A., Kessler, B.M., Reverter, D., Cook, J., Kolli, N., Gan-Erdene, T., Wilkinson, K.D., Gill, G., Lima, C.D., et al. (2004). Specific and Covalent Targeting of Conjugating and Deconjugating Enzymes of Ubiquitin-Like Proteins. Mol. Cell. Biol. *24*, 84–95.

12. Johnston, P.A. (2011). Redox cycling compounds generate H2O2 in HTS buffers containing strong reducing reagents – real hits or promiscuous artifacts? Curr. Opin. Chem. Biol. *15*, 174–182.

13. Komander, D., Clague, M.J., and Urbé, S. (2009). Breaking the chains: structure and function of the deubiquitinases. Nat. Rev. Mol. Cell Biol. *10*, 550–563.

14. Kumar, A., and Zhang, K.Y.J. (2013). Computational Investigation of SENP:SUMO Protein-Protein Interaction for Structure Based Drug Design. Mol. Inform. *32*, 267–280.

15. Kumar, A., Ito, A., Takemoto, M., Yoshida, M., and Zhang, K.Y.J. (2014). Identification of 1,2,5-Oxadiazoles as a New Class of SENP2 Inhibitors Using Structure Based Virtual Screening. J. Chem. Inf. Model. *54*, 870–880.

16. Madu, I.G., Namanja, A.T., Su, Y., Wong, S., Li, Y.-J., and Chen, Y. (2013). Identification and Characterization of a New Chemotype of Noncovalent SENP Inhibitors. ACS Chem. Biol. *8*, 1435–1441.

17. Qiao, Z., Wang, W., Wang, L., Wen, D., Zhao, Y., Wang, Q., Meng, Q., Chen, G., Wu, Y., and Zhou, H. (2011). Design, synthesis, and biological evaluation of benzodiazepine-based SUMO-specific protease 1 inhibitors. Bioorg. Med. Chem. Lett. *21*, 6389–6392.

18. Shen, L., Tatham, M.H., Dong, C., Zagórska, A., Naismith, J.H., and Hay, R.T. (2006). SUMO protease SENP1 induces isomerization of the scissile peptide bond. Nat. Struct. Mol. Biol. *13*, 1069–1077.

19. Shi, T., Han, Y., Li, W., Zhao, Y., Liu, Y., Huang, Z., Lu, S., and Zhang, J. (2013). Exploring the Desumoylation Process of SENP1: A Study Combined MD Simulations with QM/MM Calculations on SENP1-SUMO1-RanGAP1. J. Chem. Inf. Model. *53*, 2360–2368.

20. Uno, M., Koma, Y., Ban, H.S., and Nakamura, H. (2012). Discovery of 1-[4-(N-benzylamino)phenyl]-3-phenylurea derivatives as non-peptidic selective SUMO-sentrin specific protease (SENP)1 inhibitors. Bioorg. Med. Chem. Lett. *22*, 5169–5173.

21. Wen, D., Xu, Z., Xia, L., Liu, X., Tu, Y., Lei, H., Wang, W., Wang, T., Song, L., Ma, C., et al. (2014). Important Role of SUMOylation of Spliceosome Factors in Prostate Cancer Cells. J. Proteome Res. *13*, 3571–3582.

22. Wu, J., Lei, H., Zhang, J., Chen, X., Tang, C., Wang, W., Xu, H., Xiao, W., Gu, W., Wu, Y., et al. (2016). Momordin Ic, a new natural SENP1 inhibitor, inhibits prostate cancer cell proliferation. Oncotarget *5*.

23. Zhang, J.-H., Chung, T.D.Y., and Oldenburg, K.R. (1999). A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J. Biomol. Screen. *4*, 67–73.

24. Zhao, Y., Wang, Z., Zhang, J., and Zhou, H. (2016). Identification of SENP1 inhibitors through in silico screening and rational drug design. Eur. J. Med. Chem. *122*, 178–184.

25. Zorn, J.A., and Wells, J.A. (2010). Turning enzymes ON with small molecules. Nat. Chem. Biol. *6*, 179–188.

Chapter 4

Nucleotide-mediated inhibition of SENP1

1. Introduction

At the end of chapter three it was mentioned that both the class 1 and class 2 hits from HTS are known P2X1 receptor inhibitors. P2X1 receptors are a member of a family of purinergic receptors, which bind extracellular nucleotides as a signaling molecule. Extracellular nucleotide release can be induced in a number of ways including via cellular shear stresses and damages, pathogen-associated molecular patterns (PAMPs), excitatory exocytosis from neurons as a neurotransmitter, and actively by ATP-binding cassette transporters (Gorini et al., 2013; Novak, 2003). There are two types of purinergic (P2) receptors; there are the ionotropic, fast-acting P2X receptors that act as ligand gated ion channels, and there are the metabotropic P2Y receptors that function as G-protein coupled receptors (GPCRs). While P2Y receptor activation can occur via ATP, ADP, UTP, UDP, or UDP-glucose nucleotides, P2X receptors are activated only by ATP, among naturally occurring nucleotides.

‡ Carrie Johnson performed selective inhibition by ATPyS, as well as nucleotide structure activity relationships, and the kinetic characterization of ATP inhibition. In collaboration with Dr. Jorge Iñiguez-Lluhí, Carrie M. Johnson performed MANT-ATP fluorimetry. Dr. Marcelo Murai performed all NMR experiments, data analysis, and figure generation.

There are seven different subtypes of P2X receptors ($P2X_{1-7}$) in eukaryotes. Each monomer has two transmembrane domains with a large extracellular loop, but they assemble as homo- or heteromeric trimers. They are found throughout the body, including in smooth muscle, the heart, central and peripheral nerves, and immune cells and platelets. Upon receptor activation, channel opening allows for an influx of sodium and calcium ions for cell depolarization or cytosolic signaling. The receptors vary in their rates of desensitization, ranging from milliseconds to seconds, and other than common ATP binding function, P2X receptors also differ markedly in their allosteric regulation by metals, small molecules, and protons (Coddou et al., 2011a; King et al., 1996; Lorca et al., 2005; Roberts et al., 2006). With no typical structural ATP binding motifs or sequences, (*e.g*. Walker motif), found in P2X receptors, agonists and antagonists with specificity towards certain subtypes have been difficult to generate (Roberts et al., 2006; Walker et al., 1982). Based upon mutagenesis studies and recent crystal structures of zebrafish P2X4 receptors in the open and closed states, ATP binding occurs at the interface between each monomer of the trimeric receptor (Evans, 2010; Hattori and Gouaux, 2012; Hausmann et al., 2015; Kawate et al., 2009). ATP binding occurs in the bulky part of the extracellular domain in a pocket lined with numerous positively charged residues, propagating allosteric changes to the transmembrane helices to widen and open up the ion channel (Hattori and Gouaux, 2012). For receptor activation, at least two of the three ATP binding sites must be occupied (Coddou et al., 2011b). With this newer structural information, the field has been able to move beyond the existing fairly promiscuous chemical probes against P2X receptors and generate small molecule antagonists more selective to P2X receptor subtypes. Efforts targeting specific P2X receptor subtypes have been directed towards

P2X4 and P2X7 receptors for their potential therapeutic roles in chronic and neuropathic pain, inflammation, and neurodegenerative and autoimmune diseases (Bartlett et al., 2014; Jacobson and Gao, 2016). Just as it is essential to explore the binding mechanisms and breadth of ligands for modulating P2X receptor activity to generate potential therapeutics, it is essential to study potential inhibitors of SENP enzymes due to their roles in disease.

In this work, the natural agonist for P2X receptors, ATP, is studied as a potential SENP1 inhibitor. With our robust FRET assay, the mechanism of inhibition was determined in an unbiased fitting to reveal selectivity of inhibition for SENP1 over SENP2. Evidence for direct binding of ATP by SENP1 was obtained through multiple approaches. SENP1 residues responsible for selectivity were assessed and brief structure activity relationships were obtained regarding ATP analogs.

2. Materials and Methods

SENP1 and SENP2 construction, expression, and purification. Plasmids for lac promoter driven expression of N-terminal His-tagged SENP1 (AA 418-C) and SENP2 (AA 363-C) catalytic domains were constructed in the pHT2 backbone. Rosetta 2 (DE3) *E. coli* cells were transformed with the plasmids and grown under carbenicillin selection. For expression, cultures (1L per flask) were inoculated and grown at 37 °C. Upon reaching a density of OD_{600} 1.1, cultures were shifted to 18 °C and isopropyl β-D-1-thiogalactopyranoside was added to a final concentration of 0.2 mM. Cultures were harvested after a further 20 hour incubation. All subsequent purification steps were performed at 4 °C. Cells were harvested by centrifugation and suspended in Lysis Buffer

(10 mM Tris HCl pH 8.0, 500 mM NaCl, 5% glycerol, 2 mM imidazole, 1 mM β-Mercaptoethanol, 0.5 mM PMSF, 0.5 μL benzonase). Cells were disrupted by three passes through a French Press and a cleared lysate was obtained by centrifugation at 40,000 x g for 45 min at 4°C. Protein was purified from the lysate using Ni-NTA affinity chromatography and stored in 300 mM NaCl, 50 mM Tris HCl pH 7.5, and 1 mM dithiothreitol.

***In Vitro* FRET assay**. Reactions were 100 μL final and performed in high or low salt buffer solution (20 mM TrisCl pH 8.0, 250 μM Na·EDTA, 20 mM NaCl or 100 mM NaCl, 10 mM β-Mercaptoethanol, 0.1 mg mL^{-1} bovine serum albumin, 0.01% NP-40). Reactions were initiated by the addition of substrate and fluorescence was read every minute on a Spectramax M5 plate reader excitation at 405 nm and emission measured at 475 and 525 nm.

MANT-ATP Fluorimetry. Methods and data fitting in possession of Dr. Jorge Iñiguez-Lluhí.

Cloning, expression and purification

For NMR studies, the gene sequence of SENP1 (aa 419-644) containing the point mutation C603S was synthesized by GeneArt Gene Synthesis (Invitrogen) and subcloned into pET24a (Novagen) digested with EcoRI and BamHI restriction sites. All clones were verified by analytical digestion and automated sequencing. The ^{15}N-labeled SENP1 protein was expressed by growing bacterial cells in isotopically enriched M9 minimal

media. SENP1 was expressed in BL21-codon plus (DE3) and purified using affinity chromatography Ni-NTA (Qiagen) followed by SP Sepharose FF (GE Healthcare). The following point mutations were generated using site-directed mutagenesis (R449E, W465A, E469K, T495A, T503A, Q507A, T451R, K455G, K515G, and the double mutants L530R,G531K and K455G,K515G) for nucleotide binding studies. Expression and purification of mutants for enzymatic assays were carried out using one-step purification with nickel affinity beds.

NMR spectroscopy

For NMR experiments, the ^{15}N-labeled SENP1 C603S sample was prepared at a final concentration of 80-100 μM in 20 mM sodium phosphate, pH 6.8, 5 mM DTT, 0.02% NaN_3 and 10% D_2O. Chemical shift assignments of SENP1 C603S were obtained as previously described (BMRB entry19083). NMR measurements were performed using a Bruker Advance III 600-MHz spectrometer equipped with 5-mm TCI cryogenic probe. To determine binding affinity, ^{15}N-labeled SENP1 C603S was titrated with MANT-ATP, and chemical shift perturbations were measured from heteronuclear single quantum coherence (HSQC) spectra.

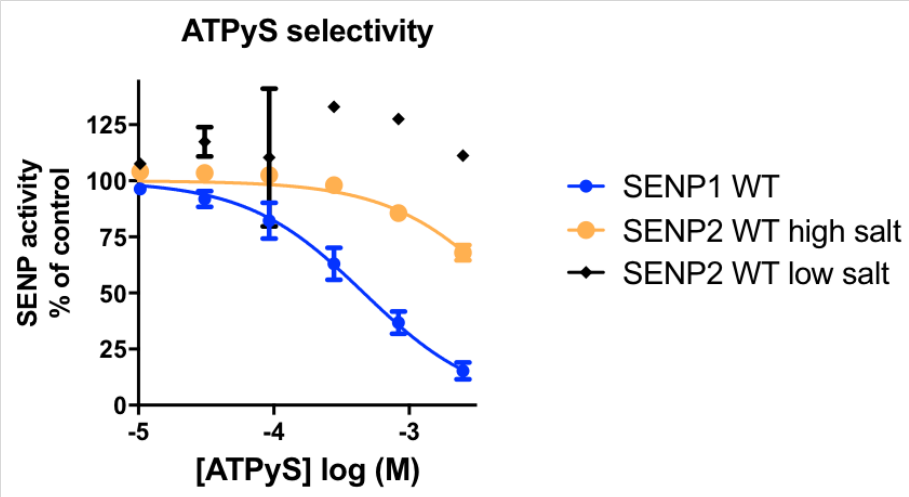
3. Results

Inhibition by ATP

Upon observation that known P2X1 inhibitors also inhibited SENP1 with some parallel pharmacology, the natural P2X1 endogenous agonist, ATP, was tested for its

effects on SENP1. Using Adenosine 5'-*O*-(3-thiotriphosphate) (ATPγS), a hydrolysis-resistant form of ATP where the gamma or tertiary phosphate contains a thio-linkage, SENP1 was inhibited in a dose-responsive manner with an IC_{50} of 463 μM (Figure 4-1A). To assess for selectivity of ATP as an inhibitor, SENP2 was also tested. Under the same low salt conditions (20 mM NaCl) that were used when testing SENP1, confounding salt effects from titration of ATPγS, a tetralithium salt, were observed so SENP2 inhibition had to be performed under its own optimal conditions (100 mM NaCl) for comparison. SENP1

A)



B)

	ATPγS IC_{50} (mM)
SENP1 WT	0.46
SENP2 WT (high salt/100 mM NaCl)	5.43

Figure 4-1 Selective inhibition of SENP1 by ATPγS.
Dose-response curves for SENP1 WT and SENP2 WT at 20 mM salt (low salt) and SENP2 WT at high salt (100 mM). B) IC_{50} values and inhibitory constants for inhibition by ATPγS of SENP1 and SENP2 at their ideal salt concentrations.

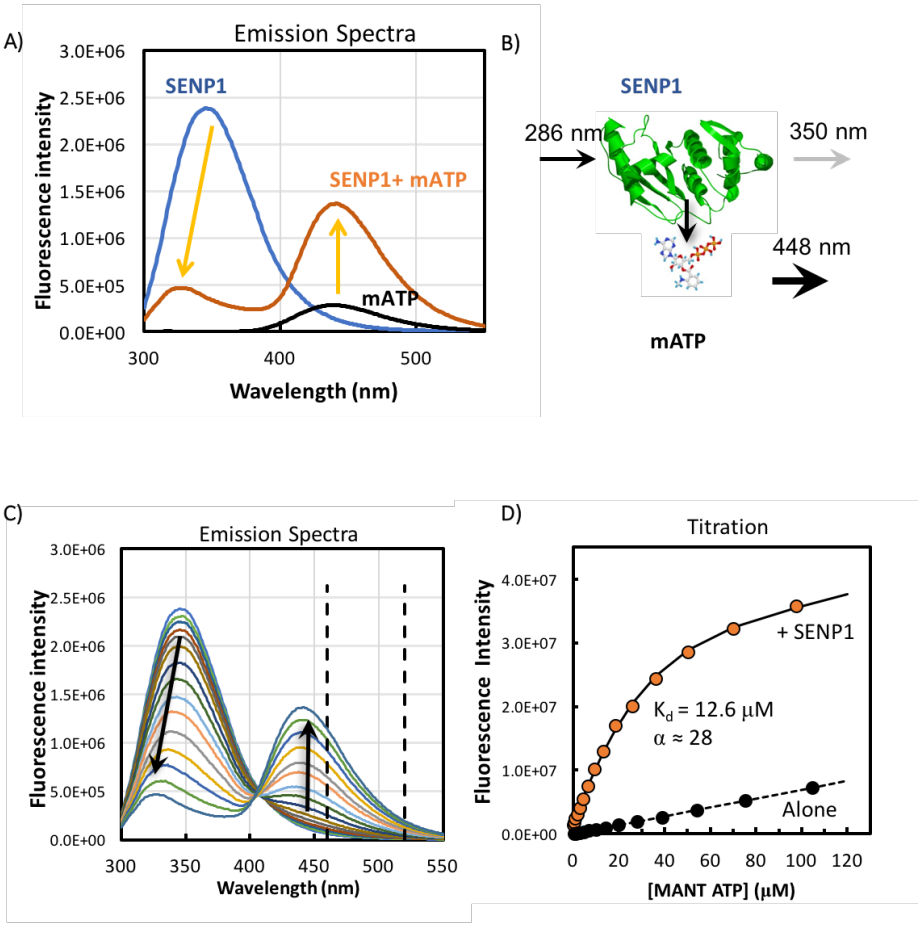
showed nearly ten-fold selectivity over SENP2 for ATP-mediated inhibition in the dose-response FRET-based assay. The estimated inhibitory constant using the Cheng-Prusoff equation reveals a nearly 20-fold ATP selectivity for SENP1 (Figure 4-1B and 4-5B).

Taking advantage of our previous detailed kinetic analysis on SENP1 and SENP2 in chapter two, the ATP inhibitory constant and mechanism of inhibition for both enzymes were obtained. With each enzyme tested under its respective optimal salt conditions, the affinity of SENP2 for SUMO1 compared to SENP1 is over 40-fold weaker (Figure 4-1B). This highlights the importance of how IC_{50} values are assay specific and are a reflection of a composite of kinetic constants that make it difficult to make strong comparisons. Therefore, while ATP is still a competitive yet weaker inhibitor of SENP2 than SENP1, it exhibits a greater degree of selectivity in the ATP binding constant than was anticipated based upon the IC_{50} values from the dose-response assay.

Direct binding using MANT-ATP

Though the FRET-based assays to measure SENP activity are quite powerful and high throughput in identifying and characterizing inhibitors, they are activity-based assays that offer indirect evidence of ATP inhibitor-SENP1 binding. To gain orthogonal evidence of ATP-mediated inhibition, an ATP analog with a fluorophore conjugated to the 2' or 3' position of the ribose ring was employed. This analog, 2'/3'-*O*-(*N*-Methyl-anthraniloyl)-ATP, abbreviated MANT-ATP, has intrinsic fluorescence with optimal excitation near 350 nm and emission at 448 nm. This optimal MANT-ATP excitation overlaps with the emission of amino acids with intrinsic fluorescent properties, like tryptophan and tyrosine, when they are excited at 280 nm. The large spectral overlap allows energy transfer to occur

between the protein and bound MANT nucleotides and resulting in a decrease in protein fluorescence at 350 nm and generation of a FRET signal at 448 nm (Figure 4-2A and B). With energy transfer as a function of distance between the fluorescent amino acids and the



	K_d (µM)
MANT ATP	12
MANT dATP	7
MANT ADP	80

Figure 4-2 Direct ATP binding via MANT–ATP.
A) Emission spectra of SENP1 alone, MANT-ATP alone, and the two together after excitation at 286 nM. B) Cartoon representation of SENP1 protein to MANT-ATP FRET activity. C) Overlay of fluorescence emission curves starting with SENP1 protein alone and titration of MANT-ATP in to the mixture. D) Representative plot of the change in fluorescence intensity at 448 nM with respect MANT-ATP concentration. E) Table of experimental dissociation constants for three nucleotide derivatives of MANT-ATP.

MANT that decays to the sixth power as distance increases, this experimental approach provides strong supporting evidence of direct ATP binding.

Upon excitation at 286 nm, titration of MANT-conjugated ATP derivatives generated a decrease in signal at 350 nm and simultaneous increase in signal at 448 nm (Figure 4-2C). In plotting the change in fluorescence intensity and quantitating the dissociation constant for the titrated MANT species, inner filter effects were accounted for to limit the pitfalls of numerous potential confounding factors (Niranjan et al., 2013). Primary inner-filter effects occur when there is non-homogenous excitation throughout the sample due to absorption of the excitation light causing each successive layer of solution to receive less and less light flux (Fonin et al., 2014; Kubista et al., 1994). Secondary inner-filter effects occur when there is reabsorption of the fluorescence, like when unbound MANT derivatives absorb the protein fluorescence. As seen in the emission spectra of SENP1 alone (blue) and MANT-ATP alone (black) after excitation at 286 nm, the protein contributes some fluorescent signal at 448 nm and the MANT derivative is capable of excitation and emission in the absence of its FRET partner (Figure 4-2A). Because MANT-ATP also partially absorbs the excitation light at 280 nm, it contributes to a strong primary inner-filter effect. Titrations of protein alone and MANT-derivatives alone were used to offset the inner filter effect contributions in calculating dissociation constants. An example of a data set used in determining the dissociation constant for MANT-ATP is shown in Figure 4-2D with a summary of three commercially available MANT-ATP derivatives shown in Figure 4-2E. In addition to providing evidence supporting direct binding of ATP with SENP1, this small data set shows a preference by SENP1 for triphosphates and loss of the electron donating effects of a hydroxyl group bonded to the ribose sugar.

Nucleotide Structure Activity Relationships

With multiple approaches supporting inhibition of SENP1 by ATP, numerous nucleotide analogs were purchased to assess structure activity relationships (SAR) in SENP1 inhibition using the FRET-based SUMO assay. A full table of the log IC_{50} values (pIC_{50}) for each nucleotide analog is found in the appendix, but notable derivatives and their trends are shown in Figure 4-3. The natural nucleotides and their deoxy- and dideoxyribose forms in Figure 4-3A, show a trend amongst all bases that the deoxyribose is more potent than ribose and the dideoxyribose form is the most potent. Additionally, the purines, adenosine and guanine, are more potent than the pyrimidines.

The effects of other ribose modifications are shown in Figure 4-3B with the potencies of the ATP, dATP, and ddATP marked for reference. In the top row of analogs, the 2' hydroxyl group off the ribose was replaced with an electron withdrawing halogen or azido group. Despite the change from electron donating (hydroxyl) to withdrawing, the potency is largely unchanged from dATP. Additionally, the size of modification in this position also seems to have little effect on potency, as atomic and ionic radii of iodine are approximately two times the size of those for fluorine. Looking at the electron donating modifications in the bottom row, the primary amine analogs have reduced potency against SENP1 whereas the methoxy analogs show similar or increased potency compared to dATP. Though both groups are of similar electron donating potential, they have opposite SAR regarding rank order effects on modification at the 2' and 3' positions. Together, these two donating groups generate inhibitors that are both more and less potent than the electron withdrawing analogs, thereby limiting a general conclusion on the roles of electronics on potency.

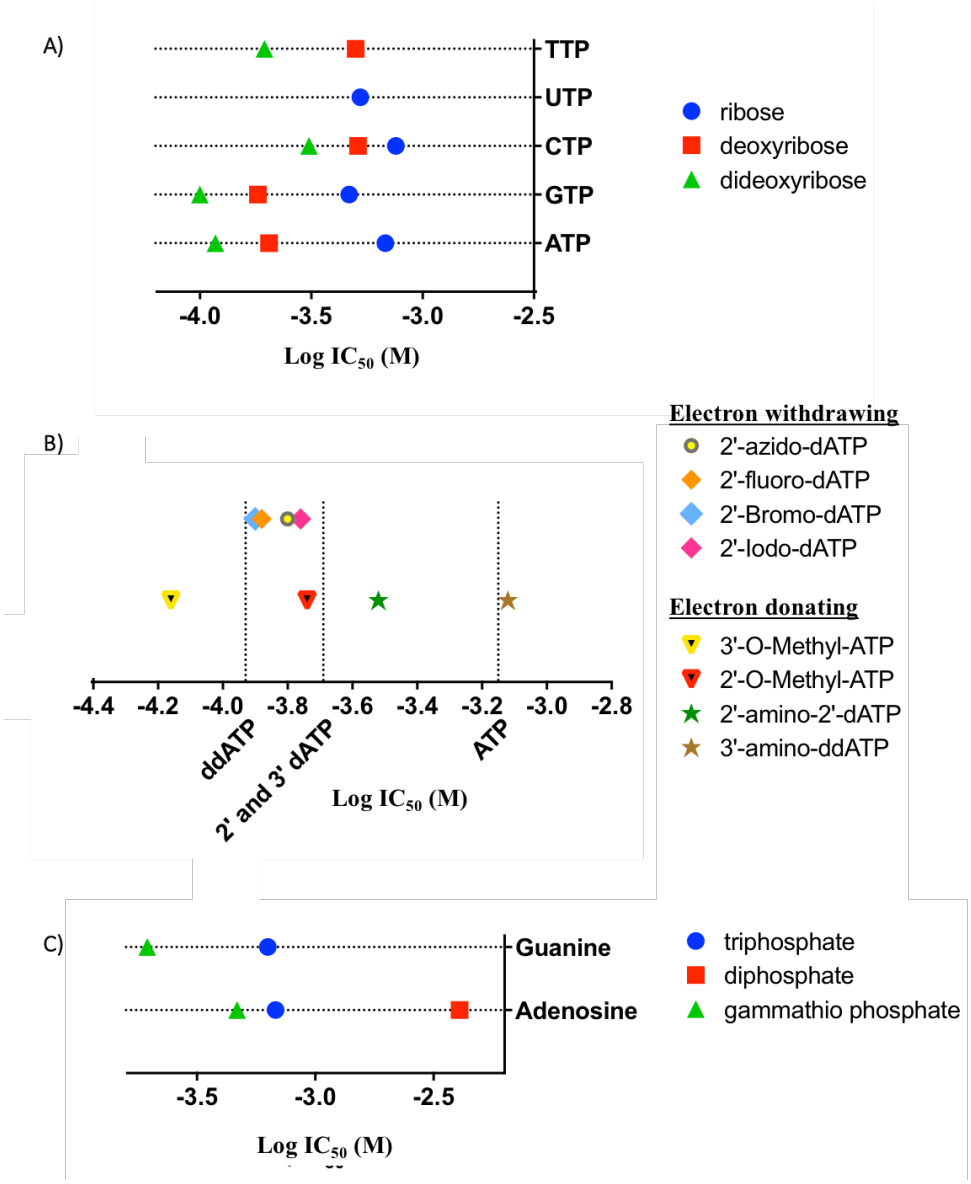


Figure 4-3 Nucleotide Structure Activity Relationships (SAR).
A) Plot of the log of the IC_{50} values of each type of nucleotide base in their ribose, deoxyribose, and dideoxyribose forms as determined by a dose-response assay using the SFCypet-SUMO1-SFYpet substrate. B) The logIC_{50} of ATP analogs containing different substituents off the ribose ring. Analogs with electron withdrawing substituents are in the top row of data points and analogs with electron donating substituents are in the lower row. The pIC_{50} values for ATP, dATP, and ddATP from (A) are noted in the x-axis with a vertical dashed line for reference. C) Effect of phosphate length in purine nucleotides on potency towards SENP1 inhibition. Adenosine monophosphate was also tested and showed no inhibitory activity.

In the last series of nucleotide SAR (Figure 4-3C), the effects of the length of the phosphate are explored. For both ATP and GTP, the hydrolysis-resistant gamma-thio analogs showed enhanced potency. The diphosphate showed a further reduction in potency and the monophosphate (data not shown) had no inhibitory effect on SENP1. Given the trend that loss of phosphate reduces inhibitory activity, it is difficult to discern if the intrinsic chemistry of sulfur in the gamma-thio analogs is responsible for the enhanced potency or if it is a product of the reduced hydrolysis of the gamma phosphate and a preservation of potency. The data from the SAR work on ATP analogs in these dose-response FRET-based experiments is in agreement with the brief SAR from the MANT-ATP fluorimetry work, where the dATP derivative had a stronger affinity than ATP and the diphosphate was weaker than both the triphosphate ATPs.

HSQC NMR analysis of nucleotide binding to SENP1

With evidence for direct ATP binding, 2D HSQC NMR was employed by a member of the lab (Dr. Marcelo Murai) to discover and evaluate the SENP1 residues likely involved in ATP binding. The same SENP1 construct was generated that was used in the previously assigned HSQC NMR of SENP1 and grown in ^{15}N labeled media (Madu et al., 2013).

Though the amide backbone of all residues are not assigned, among those that were identified, chemical shift perturbations were monitored throughout the titration of MANT-ATP into SENP (Figure 4-4A and B). The residues showing shifts in their NMR signature were mapped to the crystal structure of SENP1 according to the magnitude of their shift (Figure 4-4C). The residues in red had the largest perturbations in their amide backbone,

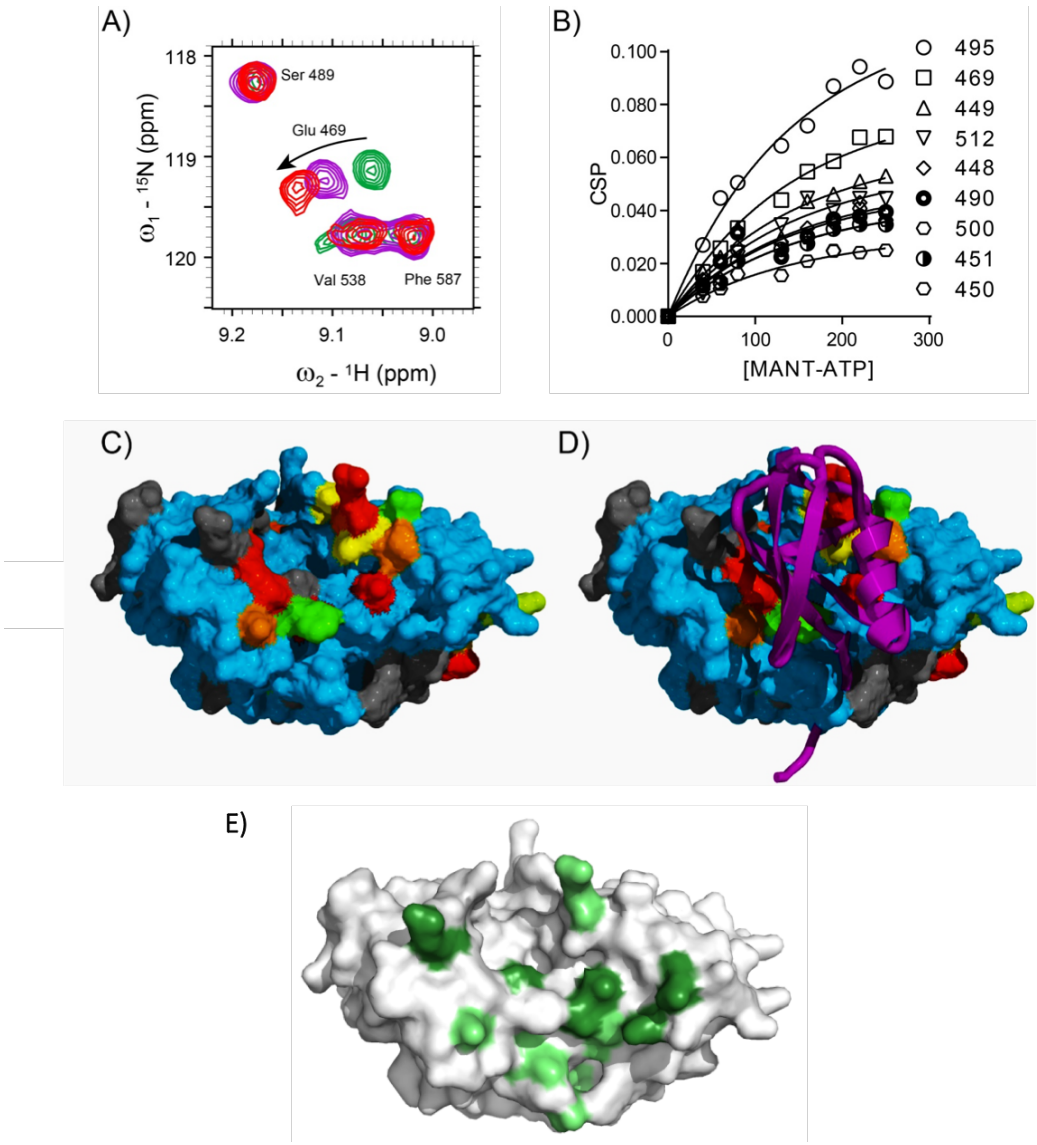


Figure 4-4 SENP1 HSQC NMR of MANT–ATP titration.
A) A portion of the HSQC NMR spectra for MANT-ATP titration into SENP1 WT, highlighting the chemical shift perturbations for the backbone amide of residue Glu 469. B) Plot of most perturbed residues in SENP1 WT as a function of MANT-ATP concentration. C) Heat map of SENP1 crystal structure depicting location and general scaling of chemical shift perturbations from MANT-ATP titration (in order from most to least perturbed: red, orange, yellow, green; blue means unperturbed and gray is for unassigned residues. D) Figure of panel C in complex with SUMO1 to give reference to the location of the most perturbed residues. E) Residues in green are those that are directly involved with SUMO binding. HSQC NMR experiments and analysis performed by Dr. Marcelo Murai.

followed by orange and yellow, and show that many of the residues most perturbed in relation to MANT-ATP binding are near the SUMO binding site where the SUMO1 the β

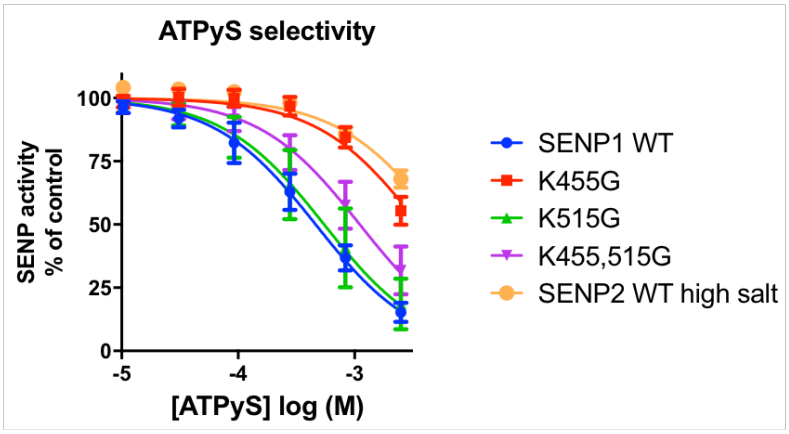
grasp domain interacts with SENP1 (Figure 4-4D). Among the residues assigned and shown, the strongest shifts in their backbone N-H bond are Thr495, Glu469, and Arg449, with Glu469 and Arg449 as amino acids known for making direct contacts with SUMO (Figure 4-4E). This overlapping location of SUMO and ATP binding supports earlier observations that ATP is a competitive inhibitor of SENP1.

Probing of SENP1 residues to understand ATP selectivity

Following analysis of the HSQC NMR spectra in the presence of ATP, a subset of the residues showing the strongest chemical shift perturbations were selected as targets for mutagenesis. In testing the mutants for general SUMO protease activity, SENP1 E469K was subjected to further testing to probe the residues responsible for the selectivity seen between SENP1 and SENP2 in inhibition by ATP. Additionally, the SENP1 mutants generated for previous studies in chapters two and three, K455G, K515G, and the double lysine mutant, K455,515G, were also tested for inhibition by ATP.

Inhibition by ATP was observed for the lysine mutants, with potencies falling in between the wild types for SENP1 and SENP2 (Figure 4-5A). Using the average IC_{50} values for ATP inhibition and the kinetic constants for each enzyme determined in chapter two, an estimated inhibitory constant was calculated using the Cheng-Prusoff equation described in Figure 3-7 of chapter three. With the estimated K_i values as a guide, detailed kinetic characterization of the WT and mutant enzymes was undertaken to generate an experimental K_i (Figure 4-5B). With the exception of SENP2 WT, the K_i values from the Cheng-Prusoff were reasonable estimations for the experimentally determined values. SENP1 WT remains the enzyme form for which ATP is most potent in inhibition, with the

A)



B)

SENP type	SENP1 (20 mM NaCl) WT	K455G	K515G	K455,515G	E469K	SENP2 (100 mM NaCl) WT
k_{cat} (s^{-1})	1.12 (0.14)	0.62 (0.19)	1.99 (0.02)	1.22 (0.03)	0.93 (0.06)	0.69 (0.17)
K_s (nM)	10.4 (2.9)	4.1 (2.8)	240 (3.6)	46.6 (1.2)	7.7 (3.8)	17.2 (3.5)
K_p (nM)	6.4 (2.3)	3.3 (3.3)	79.6 (2.6)	19.6 (0.8)	7.1 (10.3)	12.3 (3.1)
Avg. ATPγS IC_{50} (μM)	474	4250	513	1100	ND	5430
Cheng-Prusoff K_i (μM)	44.8	167	361	350	ND	797
Global Fit K_i (μM)	43.6	45.4	175	226	69.4	1920

Figure 4-5 ATP binding characterization.
A) Dose-response plot of ATPγS inhibition of different SENP1 mutants and SENP1 and SENP2 WT. B) Table of kinetic constants obtained using previous detailed kinetic analysis. Using the IC_{50} values determined via the ratiometric FRET assay, inhibitor constants were estimated using the Cheng-Prusoff (C-P) equation. Experimental inhibitory constants were obtained directly through the channel2 mediated analysis of the FRET-based substrate assay. Graphical presentation of experimental K_i values for each of the SENP1 mutants and SENP2 WT. ND = no data.

K515G and K455,515G showing the largest deviations from SENP1 WT. For K455G, the lysine at this position enhances substrate binding but has no effect on the ATP inhibitory constant, suggesting that there are non-overlapping residues important for ATP versus

SUMO binding. In contrast, K515G shows that this residue is important for both ATP and SUMO substrate binding, while the double mutant data shows that the addition of the K455G mutation is not sufficient to overcome the loss of binding affinity created by K515G. Despite strong chemical shift perturbations, change in charge at 469 from Glu to Lys has little to no effect on substrate binding and ATP-mediated inhibition.

4. Discussion

This work exhibits a unique transition of leveraging the parallels of small molecule inhibitors of an unrelated protein to the discovery of novel inhibitors for SENP1. In probing ATP-mediated inhibition through orthogonal approaches, regions of SENP1 and specific residues within those regions were identified for their involvement in ATP binding. Additionally, preliminary structure activity relationships offer insight into potential further rational drug design for development of selective, small molecule probes to then selectively study SENPs in more biological settings than tests with purified protein.

Unlike other identified SENP inhibitors, the structural evidence for ATP-mediated inhibition suggests that it binds away from the catalytic site and instead in the SENP1 pocket responsible for binding to the globular portion of SUMO, the β grasp domain. As seen in the conserved binding site for ATP in P2X receptors, one side of this binding pocket on the surface of SENP1 contains a patch of positively charged residues that could be responsible to binding to the ATP triphosphate moiety. Some of these positively charged residues are the same residues or near the same residues most perturbed in the HSQC NMR experiments with MANT-ATP. For labs who employed *in silico* screening or structure-based drug design to initiate SENP inhibition and to further advance potency using docking

studies, inhibition was targeted near the catalytic site (Chen et al., 2012; Madu et al., 2013). Specificity for SENP1 over other cysteine proteases and within the SENP family was attempted by incorporating the steric constraints of the tryptophan tunnel and designing aromatic portions to engage in pi-bond stacking with His529 and Phe496. While a desirable strategy for selectivity, it could make potent inhibition difficult. Most of the binding affinity by SENPs is driven by interaction with SUMO's β-grasp domain and the SUMO C-terminal tail seems to sway in both SUMO-bound and unbound states (Chen et al., 2014). Whether substrate bound or not, conformational dynamics show that the catalytic tunnel experiences extensive dynamics, opening and closing over a range of timescales. The tryptophan residue that serves as the lid for the catalytic tunnel generation, Trp465, is especially susceptible to these conformational changes. When SUMO and SENP1 bind through SUMO's β-grasp domain, it induces an allosteric effect that propagates from this β-grasp binding surface to the catalytic tunnel to align the catalytic residues for optimal proteolysis. Therefore, if a potential inhibitor binds to the minimal contacts near the catalytic tunnel, binding of the globular portion of SUMO and the robust allostery it induces could kick off the inhibitor that bound to the enzyme conformation where SUMO is not also bound. Alternatively, using strategies to target a small molecule inhibitor near the SENP1 catalytic tunnel in the SUMO-unbound state, it could operate as a negative allosteric modulator, locking the residues in a confirmation suboptimal for catalysis.

While much is known about the enzymology and SUMO isoform selectivity among the SENP family enzymes, very little is known about their regulation. Considering their roles in precursor processing to allow SUMOylation to occur up front and then in reversing

the modification, it is essential to gain a better understanding of how SENP activity is regulated. There is some understanding of SENP regulation as a product of their variable N-terminal domains (NTDs) and role in dictating localization. In addition, the NTDs of many of the SENPs contain SUMO-interacting motifs (SIMs) described in chapter one that may play a role in regulating their activity, but preliminary work suggests that they are not relevant for enzyme activity in vitro (Kerscher, 2007; Mendes et al., 2016). As an isolated observation, SENP1 has been found to be SUMOylated itself by SUMO2/3 but the ramifications of that modification on SENP1 function are unknown (Sharma et al., 2013). Just as there are a few proteins involved in the SUMOylation compared to the ubiquitination system, there is also a similar discrepancy in the number of proteins involved in their deconjugation systems. There are nearly 100 deubiquitinases but only six deSUMOylases, yet both PTMs are exquisitely involved in many essential cell functions and cause disease or death when they go awry. Clearly, regulation of the SUMO system in some manner is necessary for homeostasis.

With ATP as an abundant biological molecule that is at the crux of cell metabolism and function, and SUMO's role in processes regarding mitosis, DNA repair, and cell stress, it is possible that SENP activity could be linked to energy regulation in cells. Though the inhibitory and binding constants for ATP are not potent relative to most relevant small molecule therapeutics, the values are in this case in the context of intracellular ATP concentrations, which range approximately from low micromolar to 5 mM (Ando et al., 2012; Imamura et al., 2009; Novak, 2003). Throughout a cell's life cycle, ATP levels are dynamic and can oscillate depending on the cell stage as their demands and homeostatic challenges change (Tyson, 2002; da Veiga Moreira et al., 2015). Therefore, with similar

parallels in SUMO's role in responding to these various homeostatic challenges and progressions through the cell cycle, it is possible that energy levels as determined by ATP concentrations could regulate SENP activity. The logical next step might be to test the effects of manipulating ATP levels on protein SUMOylation and function of specific SUMOylation-regulated transcription factors. The challenge in this approach is that in assessing SENP function, confounding factors would make analysis difficult because ATP is necessary for initial SUMOylation. Additionally, with SENPs playing roles in both the maturation and reversal of SUMO modification, teasing out the role of ATP in the individual components of SUMO processing, conjugation, and deconjugation would be particularly challenging. Nonetheless, the potential for discovering an understanding of regulation of the SENP family of enzymes is intriguing.

In addition to the potential implications for a SENP regulatory mechanism, our initial foray into nucleotide structure-activity relationships can be leveraged to generate increasingly selective, potent, and ultimately cell permeable inhibitors. This will not be simple, as our initial attempts at generating these types of compounds with the help of Jeff Zwicker in Dr. Scott Larsen's lab were unsuccessful. Because of the commonality of ATP binding in the human proteome and the potential for cross reactivity, it is imperative to develop selective inhibitors. The purinergic receptor field has begun to be able to generate ligands selective not just for purinergic receptors in general, but for specific subtypes within the ionotropic family of purinergic receptors. These new generation ligands were tested and, with their enhanced selectivity, do not inhibit SENP1 (data not shown). Due to conflicting data across species types and receptor truncations, the purinergic field has also recognized the importance of using the human version of the protein, and the full-length

version as well. With the NTD of SENPs containing SIMs and serving functions in localization, it is important to try to use full-length enzyme when possible. Similarly, Mendes *et al*. found new results when characterizing SENPs upon using full length enzyme as opposed to the catalytic domain that was used in previous studies (Mendes et al., 2016). As the SUMO field advances, transitioning into using full-length SENP enzyme will likely become necessary to achieve a detailed understanding despite the difficulties in its synthesis and purification.

5. References

1. Ando, T., Imamura, H., Suzuki, R., Aizaki, H., Watanabe, T., Wakita, T., and Suzuki, T. (2012). Visualization and Measurement of ATP Levels in Living Cells Replicating Hepatitis C Virus Genome RNA. PLoS Pathog. *8*.

2. Bartlett, R., Stokes, L., and Sluyter, R. (2014). The P2X7 Receptor Channel: Recent Developments and the Use of P2X7 Antagonists in Models of Disease. Pharmacol. Rev. *66*, 638–675.

3. Chen, C.-H., Namanja, A.T., and Chen, Y. (2014). Conformational flexibility and changes underlying activation of the SUMO-specific protease SENP1 by remote substrate binding. Nat. Commun. *5*.

4. Chen, Y., Wen, D., Huang, Z., Huang, M., Luo, Y., Liu, B., Lu, H., Wu, Y., Peng, Y., and Zhang, J. (2012). 2-(4-Chlorophenyl)-2-oxoethyl 4-benzamidobenzoate derivatives, a novel class of SENP1 inhibitors: Virtual screening, synthesis and biological evaluation. Bioorg. Med. Chem. Lett. *22*, 6867–6870.

5. Coddou, C., Yan, Z., Obsil, T., Huidobro-Toro, J.P., and Stojilkovic, S.S. (2011a). Activation and Regulation of Purinergic P2X Receptor Channels. Pharmacol. Rev. *63*, 641–683.

6. Coddou, C., Stojilkovic, S.S., and Huidobro-Toro, J.P. (2011b). Allosteric modulation of ATP-gated P2X receptor channels. Rev. Neurosci. *22*, 335–354.

7. Evans, R.J. (2010). Structural interpretation of P2X receptor mutagenesis studies on drug action. Br. J. Pharmacol. *161*, 961–971.

8. Fonin, A.V., Sulatskaya, A.I., Kuznetsova, I.M., and Turoverov, K.K. (2014). Fluorescence of Dyes in Solutions with High Absorbance. Inner Filter Effect Correction. PLOS ONE *9*, e103878.

9. Gorini, S., Gatta, L., Pontecorvo, L., Vitiello, L., and la Sala, A. (2013). Regulation of innate immunity by extracellular nucleotides. Am. J. Blood Res. *3*, 14–28.

10. Hattori, M., and Gouaux, E. (2012). Molecular mechanism of ATP binding and ion channel activation in P2X receptors. Nature *485*, 207–212.

11. Hausmann, R., Kless, A., and Schmalzing, G. (2015). Key Sites for P2X Receptor Function and Multimerization: Overview of Mutagenesis Studies on a Structural Basis. Curr. Med. Chem. *22*, 799–818.

12. Imamura, H., Nhat, K.P.H., Togawa, H., Saito, K., Iino, R., Kato-Yamada, Y., Nagai, T., and Noji, H. (2009). Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators. Proc. Natl. Acad. Sci. *106*, 15651–15656.

13. Jacobson, K.A., and Gao, Z.-G. (2016). CHAPTER 11:Allosteric Modulators of Adenosine, P2Y and P2X Receptors. In Allosterism in Drug Discovery, pp. 247–270.

14. Kawate, T., Michel, J.C., Birdsong, W.T., and Gouaux, E. (2009). Crystal structure of the ATP-gated P2X4 ion channel in the closed state. Nature *460*, 592–598.

15. Kerscher, O. (2007). SUMO junction--what's your function? New insights through SUMO-interacting motifs. EMBO Rep. *8*, 550–555.

16. King, B.F., Ziganshina, L.E., Pintor, J., and Burnstock, G. (1996). Full sensitivity of P2times2 purinoceptor to ATP revealed by changing extracellular pH. Br. J. Pharmacol. *117*, 1371–1373.

17. Kubista, M., Sjöback, R., Eriksson, S., and Albinsson, B. (1994). Experimental correction for the inner-filter effect in fluorescence spectra. Analyst *119*, 417–419.

18. Lorca, R.A., Coddou, C., Gazitúa, M.C., Bull, P., Arredondo, C., and Huidobro-Toro, J.P. (2005). Extracellular histidine residues identify common structural determinants in the copper/zinc P2X2 receptor modulation. J. Neurochem. *95*, 499–512.

19. Madu, I.G., Namanja, A.T., Su, Y., Wong, S., Li, Y.-J., and Chen, Y. (2013). Identification and Characterization of a New Chemotype of Noncovalent SENP Inhibitors. ACS Chem. Biol. *8*, 1435–1441.

20. Mendes, A.V., Grou, C.P., Azevedo, J.E., and Pinto, M.P. (2016). Evaluation of the activity and substrate specificity of the human SENP family of SUMO proteases. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1863*, 139–147.

21. Niranjan, Y., Ungureanu, D., Hammarén, H., Sanz-Sanz, A., Westphal, A.H., Borst, J.W., Silvennoinen, O., and Hilhorst, R. (2013). Analysis of steady-state Förster resonance energy transfer data by avoiding pitfalls: Interaction of JAK2 tyrosine kinase with N-methylanthraniloyl nucleotides. Anal. Biochem. *442*, 213–222.

22. Novak, I. (2003). ATP as a Signaling Molecule: the Exocrine Focus. Physiology *18*, 12–17.

23. Roberts, J.A., Vial, C., Digby, H.R., Agboh, K.C., Wen, H., Atterbury-Thomas, A., and Evans, R.J. (2006). Molecular properties of P2X receptors. Pflüg. Arch. *452*, 486–500.

24. Sharma, P., Yamada, S., Lualdi, M., Dasso, M., and Kuehn, M.R. (2013). Senp1 Is Essential for Desumoylating Sumo1-Modified Proteins but Dispensable for Sumo2 and Sumo3 Deconjugation in the Mouse Embryo. Cell Rep. *3*, 1640–1650.

25. Tyson, J.J. (2002). Biochemical Oscillations. In Computational Cell Biology, (Springer, New York, NY), pp. 230–260.

26. da Veiga Moreira, J., Peres, S., Steyaert, J.-M., Bigan, E., Paulevé, L., Nogueira, M.L., and Schwartz, L. (2015). Cell cycle progression is regulated by intertwined redox oscillators. Theor. Biol. Med. Model. *12*, 10.

27. Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. *1*, 945–951.

Chapter 5

Conclusions and future directions

1. Overview

Since the discovery of SUMO as a PTM in 1996, many have sought to understand its targets, its roles in modulating protein function, and the implications of its effects in disease contexts. Though many SUMO targets have been found in the nucleus, SUMOylated proteins have also been found in the cytoplasm and in transmembrane proteins of the endoplasmic reticulum, plasma membrane, Golgi apparatus, and outer mitochondrial membrane (Flotho and Melchior, 2013). As a reversible modification with highly dynamic and rapid cycles of SUMOylation and deSUMOylation, SUMO can act as a molecular switch in modulating protein function. With roles in numerous cancers, Huntington's Disease, Alzheimer's Disease, amyotrophic lateral sclerosis, Parkinson's Disease, dilated cardiomyopathy, and other diseases, numerous studies of the SUMO system have been performed (Sarge and Park-Sarge, 2011). In the context of prostate cancer, preservation of SUMOylation of the androgen receptor (AR) helps limit transcription of AR-regulated genes and therefore the proliferation of the cancer. As described in chapter 1, antagonizing this suppression of transcription is the SUMO protease SENP1. SENP1 is overexpressed in advanced prostate cancer and can initiate progression of healthy prostatic tissue into the cancer precursor, prostatic intraepithelial neoplasia

(Bawa-Khalfe et al., 2010). Therefore, we attempted to identify novel small molecule inhibitors of SENP1.

2. Summary and impact of results

In this work, a FRET-based approach was employed to initialize the search for SENP1 inhibitors with potential therapeutic implications. A ratiometric analysis of the substrate's fluorescent emissions was ideal for small-volume high throughput screening because it can account for small changes in substrate concentration from well to well. After screening over 500,000 compounds, an iterative process of utilizing different set-ups of the ratiometric FRET assay in dose-response and time course formats generated the desired funnel-like effect of the narrowing of candidate compounds in drug discovery.

While assessing the selectivity of some of these inhibitors for SENP1 versus SENP2, a differential salt sensitivity of the enzymes was observed. Though it may seem like a minor issue, this observation is surprisingly novel to the SUMO protease field. Due to the array of diseases where SENP1 has implications, there have been many efforts to develop SENP1 inhibitors. Yet when these groups are testing inhibition against SENP2 and attempting to further drive the potency and selectivity towards SENP1, potential structure activity relationships may be misleading because of the suboptimal testing conditions. Through the careful monitoring and standardization of channel 2 signal at 525 nm, further analysis of these salt effects offered a more detailed characterization of the consequences ionic strength on SENP1 vs SENP2 proteolytic activity. This detailed characterization for teasing out the individual kinetic constants facilitated probing specific

residues for their roles in substrate binding, product binding, and catalysis. Lastly, this approach to analysis also permits unbiased determination of the mechanism of inhibition of identified small molecule inhibitors, which can aid in the development of better inhibitors. The results presented here demonstrate the importance of characterizing each enzyme separately and the robust utility of the generated FRET substrate in the discovery and further characterization of SENP enzyme inhibitors regarding their mechanism of inhibition, structure activity relationships, inhibitory constants, and residues important for inhibitor binding.

Among the most unique findings of this work and with potentially the greatest impact is the parallel pharmacology of small molecule inhibitors between SENP1 and ionotropic purinergic receptors. Observation of SENP1 inhibition by these small molecule antagonists and by ATP was certainly unexpected given no structural similarities to the two proteins and the absence of any canonical ATP binding motifs. ATP was found to be a competitive inhibitor with Lys 515 in SENP1 showing evidence that it is involved in ATP binding. Notably, this residue is in the same binding pocket of SENP1 where SUMO binds, and it is where a negatively charged surface patch of SUMO1 interacts with a positively charged patch of SENP1. As a competitive inhibitor with three negatively charged phosphate groups, the partial dependence on a SENP1 lysine residue, among others, helps in creating a more cohesive understanding. The HSQC NMR and MANT-ATP assays assessed direct binding instead of SENP activity, generating further supporting evidence for ATP-SENP1 interactions. In agreement with the activity-based results, some of the backbone residues sensitive to ATP in HSQC NMR are near Lys 515. If the R-group amine in the lysine were assigned it would be advantageous to monitor potential chemical shift

perturbations at this extended R group than just the effects on the backbone nitrogen-proton 2D signal. Though the initial structure activity relationships for nucleotides are brief due to our limited supply, they provide a potential starting point for guiding the further development of more potent and biologically stable SENP1 inhibitors.

3. Regulation of the SUMO and AR systems

There are a multitude of reports on the functional effects of AR SUMOylation and the extensive biochemical, cellular, and clinical validation of SENP1 as a protein potentiating AR signaling and tumor growth. Yet despite the plethora of SUMO literature in the past two decades, there remains very little understanding as to the exact mechanism of how SUMOylation represses AR transcription. AR is predominantly deSUMOylated in the cytosol and becomes SUMOylated upon ligand binding-induced activation and translocation to the nucleus (Kaikkonen et al., 2008). There is evidence for SUMO-induced chromatin level effects, including the recruitment of histone deacetylases (co-repressors) to condense chromatin and prevent DNA-binding access by the general transcriptional assembly (Girdwood et al., 2003; Yang and Sharrocks, 2004). Additionally, there is evidence for both a SUMO-mediated enhancement and reduction in AR occupancy at DNA androgen response elements, possibly in a gene selective manner (Rytinki et al., 2012). Pivoting within the DNA occupancy idea, hypotheses related to receptor clearance from the DNA response element and cycling of rounds of receptor have been argued both for and against suppressing transcriptional activity, especially in the context of competition for

modification between SUMO and ubiquitin-induced degradation (Chymkowitch et al., 2011; Kodadek et al., 2006).

Beyond effects on AR, regulation of SUMOylation and deSUMOylation is also poorly understood. Under resting conditions, most SUMO1 is conjugated to a substrate and there is large pool of free SUMO2/3. Cell stress has reliably been shown to increase SUMO2/3 modification across the cellular proteome and there is supporting evidence for group-based SUMOylation where related target proteins are already in proximity to each other (Hendriks and Vertegaal, 2016; Saitoh and Hinchey, 2000). For the SENPs, comprehension of their regulation is limited to the localization-directing effects of their differing N-terminal domains. Importantly, regulation of SENP enzymes is especially integral for exquisite and timely implementation of SUMO functions due to its role in both precursor processing (SENP1 and SENP2) and deSUMOylation (true for all six SENPs). The results presented in chapters 2 and 4 point to two potential factors that could play a role in SENP1 and SENP2 regulation.

First, in chapter 2 a differential salt sensitivity between SENP1 and SENP2 was observed with respect to their activity, kinetic constants, and SUMO processing isoform preference. While the consequences of the differential sensitivity were discussed in the context of testing and developing selective small molecule inhibitors, there are also potential implications for regulation. In mammalian cells, the salt concentration in the nucleus is around twice that in the cytosol (~150 mM vs ~300 mM). SENP1 and SENP2 have been observed in both the nucleus and cytoplasm (Moore and Morrill, 1976; Terry et al., 2011). Studies across the SUMO field have revealed that dynamic regulation of SUMOylation and its functions is essential for development and numerous other signaling

pathways. While the salt-based data presented in this thesis were based on the catalytic domain of the enzymes and not the full-length enzyme, the contributions of electrostatics to SENP affinities for SUMO substrate and to peptide hydrolysis may help tune the activity of these proteases as they shuttle between different cellular locales.

Regarding chapter 4, nucleotides and nucleotide analogs were found to inhibit SENP1 and to a lesser extent, SENP2. With EDTA in the reaction buffer and no magnesium added, another unique characteristic of this ATP effect is that their inhibition is not magnesium-dependent unlike other enzymes that use ATP. As the molecule that functions as the primary energy currency in a cell, ATP is tied to the essential function of healthy cell growth and homeostasis but also to cancer cell growth. In highly proliferative cancers, the demands for ATP are substantially higher than those of healthy cells. The cancer cells utilize ATP at a high rate to maintain their nearly unabated pace of growth and metabolism. Therefore, in prostate cancer, the proliferative drive of the disease may lead to high turnover in ATP and an overall reduced cellular [ATP] as the cells strive to keep up with energy demands. In this situation, should SENP1 activity be regulated by ATP inhibition near physiological concentrations, reduced ATP levels in cancer cells could further augment the role of SENP1 in prostate cancer progression.

4. Exploration of future prostate cancer therapeutics

As described in chapter 1, beyond surgery, the three primary chemotherapeutic means to prostate cancer therapy all target the ligand-binding domain of AR either directly or indirectly. Chemical castration, AR antagonism, and direct inhibition of androgen

synthesizing enzymes all cause initial shrinkage in the size of both the healthy and cancerous prostatic tissue. When patients invariably fail these therapies, there are a few cytotoxic agents that can be employed to extend survival by a few months. In other areas of cancer treatment such as melanoma and non-small cell lung cancer, immunotherapy has recently generated considerable momentum. In prostate cancer, new generation immunotherapy has been around since the FDA approved Sipuleucel-T in 2010. In this treatment, dendritic cells are harvested from the patient and cultured in the presence of antigens from the protein prostatic acid phosphatase (PAP) and an immune-stimulant for immune cell maturation. In programming these immune cells to target PAP via co-culture, an amplified culture of these programmed cells is re-injected into the patient in hopes the immune cells will specifically recognize and ultimately kill the cancer cells. Like the cytotoxic agents, this immunotherapy has been shown to increase survival by around four months (Small et al., 2006).

Considering the state of the disease and the lack of therapeutics that are effective over periods of time longer than a few months, there is continued motivation to develop new therapies. As noted, immunotherapy is a modality that continues to develop and is offering exciting results in some other cancers despite the limited duration of efficacy thus far in prostate cancer. Though there is worthy excitement over immunotherapy potentials across cancer types, within advanced prostate cancer, the disease is still largely dependent on AR. Some efforts to curb this disease or generate cures are looking at alternative mechanisms to modulate AR-dependent transcription other than the traditional approaches related to the ligand binding domain. Looking at other AR domains, there have been some reports of small molecules to target the DNA binding domain of AR (Dalal et al., 2014; Li

et al., 2014; Lim et al., 2014). Others have targeted the N-terminal domain (NTD) of AR and its activation function (AF)-1 region to attenuate transactivation of AR and AR-dependent transcriptional output (Andersen et al., 2010). Improvements in drug delivery, selectivity, and durability for administration of siRNA allow for their future development as potential therapies targeting AR or other proteins implicated in advanced prostate cancer growth such as the ERG family of transcription factors (Feng et al., 2014; Tomlins et al., 2008). In a unique approach developed by Dr. Craig Crews, the selective degradation of AR is employed as an alternative approach. In these proteolysis-targeting chimeras (PROTACs), a heterobifunctional ligand, is used. These ligands have a target protein binding molecule on one end, dihydrotestosterone, and a protein degradation machinery recruiting unit, a ubiquitin E3 ligase, on the other end with a linker in between (Ottis and Crews, 2017).

While this sampling of strategies and others not mentioned to treat advanced prostate cancer have varying degrees of utility and prospective ability to become a therapeutic, targeting SENP1 offers unique advantages within the disease context. With SUMOylation occurring in the NTD of AR, a SENP1 inhibitor could synergize with existing therapies that target the LBD. In addition to the roles of SENP1 in potentiating AR activity and completing a positive feedback loop, other advantageous reasons for inhibiting SENP1 relate to its roles in radiosensitization, senescence, and metastasis. In a lung cancer model, inhibition of SENP1 induced radiosensitization, suggesting that it could potentially decrease the amount of radiation needed to treat cancerous tissue (Wang et al., 2013). Both inhibition of SENP1 and hyperSUMOylation each induce senescence. (Kaikkonen et al., 2008; Li et al., 2006). Lastly, SENP1 promotes epithelial to

mesenchymal transition and silencing of SENP1 induces apoptosis and slows proliferation and cell migration. These are all important steps in metastasis, which is responsible for 90% of cancer deaths (Hanahan and Weinberg, 2011). Therefore, with SENP1 inhibition offering the many advantageous outcomes described, strategies to inhibit it remain attractive.

5. Future directions

Moving forward, the importance of the full-length enzyme in assessing regulatory mechanisms of SENP activity must be recognized. Separate from the catalytic domain, the NTD is responsible for the localization of SENPs and shows less similarity between SENP enzymes. Quantitative purification of full-length enzyme that contains this NTD remains difficult and lack of crystal structures of the non-catalytic domain regions makes it difficult to hypothesize what changes they might initiate in SENP regulation or selective inhibition.

The nucleotide effects on SENP1 activity are particularly intriguing. Back in 1956, Dr. Otto Warburg observed that cancer cells have an altered metabolism that allows them to survive stressful conditions while maintaining necessary proliferation for cancer progression and metastasis (Warburg, 1956). In this altered metabolic state, the cancer cells increase their dependence on glycolysis over oxidative phosphorylation to meet their energy demands and lactic acid production rises. To accommodate these demands, these cells also increase their rate of intracellular glucose importation, a characteristic utilized for positron emission tomography (PET) imaging with fluorodeoxyglucose in cancer patients. With respect to this Warburg effect found in many cancer cells, changes in

cellular metabolism due to reduced utilization of oxidative phosphorylation and increased reliance on glycolysis may provide key support in regulating SENP1 activity. With a potential for quick turnover between ATP production and rapid utilization in cancer cells, as ATP levels drop in prostate cancer cells, SENP1 activity may become enhanced to allow for more rapid proliferation. As described above, inhibiting SENP1 would then not only repress AR transcriptional activity, but it could also lead to slowing cancer progression, radiosensitization, senescence, and lower metastatic potential.

The challenge in elucidating a potential role for ATP-mediated regulation of SENP1 lies within the entire SUMOylation pathway. The first step in preparation of conjugating mature SUMO to a substrate requires ATP. Therefore, if ATP levels were manipulated in an attempt to assess effects on SENP1 activity in the cell, the SUMOylation process itself would be impaired and confound the results. Therefore, the best way to get around this might be to develop a more selective and biologically stable ATP-mimetic to study SENP1 activity in a spatial and temporal manner.

An additional interesting future research aim could involve comparing SENP1 expression levels and ATP levels. Analysis of three of the most utilized prostate cancer cell lines has previously reported that under normoxic conditions, LNCaP cells rely on respiration over glycolysis in contrast to PC3 and DU145 cells (Higgins et al., 2009). This behavior for mechanism of ATP production parallels the concentrations of ATP within these cells. Glycolysis is much less efficient than oxidative phosphorylation, and as anticipated, the PC3 and DU145 cells have much lower ATP levels than LNCaP cells. Within the context of these cells, it would be worth seeing if SENP1 expression levels have any relation to cellular ATP levels in the individual cell lines. In regards to the current

model, lower ATP levels could lead to higher SENP1 levels due to the loss of SUMO-based restraint on AR activity and the fact that SENP1 is a direct AR gene product. Additionally, transfection with labeled SUMO or blotting for endogenous SUMO would also be worth pursuing to see if levels of SUMO conjugated substrates differ with respect to ATP levels.

Ultimately, SENP1 remains an attractive therapeutic target but a difficult chemical one. With over two decades of a head-start, the ubiquitin field has also struggled to develop selective small molecule inhibitors (Huang and Dixit, 2016). Even beyond potential therapeutic implications, the integral role of SUMO in numerous cell functions necessitates the development of better chemical tools to study the SUMO system in more dynamic, biologically relevant settings. The findings and assay development presented in this thesis serve as valuable tools in advancing this process.

6. References

1. Andersen, R.J., Mawji, N.R., Wang, J., Wang, G., Haile, S., Myung, J.-K., Watt, K., Tam, T., Yang, Y.C., Bañuelos, C.A., et al. (2010). Regression of Castrate-Recurrent Prostate Cancer by a Small-Molecule Inhibitor of the Amino-Terminus Domain of the Androgen Receptor. Cancer Cell *17*, 535–546.

2. Bawa-Khalfe, T., Cheng, J., Lin, S.-H., Ittmann, M.M., and Yeh, E.T.H. (2010). SENP1 Induces Prostatic Intraepithelial Neoplasia through Multiple Mechanisms. J. Biol. Chem. *285*, 25859–25866.

3. Chymkowitch, P., Le May, N., Charneau, P., Compe, E., and Egly, J.-M. (2011). The phosphorylation of the androgen receptor by TFIIH directs the ubiquitin/proteasome process. EMBO J. *30*, 468–479.

4. Dalal, K., Roshan-Moniri, M., Sharma, A., Li, H., Ban, F., Hessein, M., Hsing, M., Singh, K., LeBlanc, E., Dehm, S., et al. (2014). Selectively Targeting the DNA-binding Domain of the Androgen Receptor as a Prospective Therapy for Prostate Cancer. J. Biol. Chem. *289*, 26417–26429.

5. Feng, F.Y., Brenner, J.C., Hussain, M., and Chinnaiyan, A.M. (2014). Molecular Pathways: Targeting ETS Gene Fusions in Cancer. Clin. Cancer Res. *20*, 4442–4448.

6. Flotho, A., and Melchior, F. (2013). Sumoylation: A Regulatory Protein Modification in Health and Disease. Annu. Rev. Biochem. *82*, 357–385.

7. Girdwood, D., Bumpass, D., Vaughan, O.A., Thain, A., Anderson, L.A., Snowden, A.W., Garcia-Wilson, E., Perkins, N.D., and Hay, R.T. (2003). p300 Transcriptional Repression Is Mediated by SUMO Modification. Mol. Cell *11*, 1043–1054.

8. Hanahan, D., and Weinberg, R.A. (2011). Hallmarks of Cancer: The Next Generation. Cell *144*, 646–674.

9. Hendriks, I.A., and Vertegaal, A.C.O. (2016). A comprehensive compilation of SUMO proteomics. Nat. Rev. Mol. Cell Biol. *17*, 581–595.

10. Higgins, L.H., Withers, H.G., Garbens, A., Love, H.D., Magnoni, L., Hayward, S.W., and Moyes, C.D. (2009). Hypoxia and the metabolic phenotype of prostate cancer cells. Biochim. Biophys. Acta BBA - Bioenerg. *1787*, 1433–1443.

11. Huang, X., and Dixit, V.M. (2016). Drugging the undruggables: exploring the ubiquitin system for drug development. Cell Res. *26*, 484–498.

12. Kaikkonen, S., Jaaskelainen, T., Karvonen, U., Rytinki, M.M., Makkonen, H., Gioeli, D., Paschal, B.M., and Palvimo, J.J. (2008). SUMO-Specific Protease 1 (SENP1) Reverses the Hormone-Augmented SUMOylation of Androgen Receptor

and Modulates Gene Responses in Prostate Cancer Cells. Mol. Endocrinol. *23*, 292–307.

13. Kodadek, T., Sikder, D., and Nalley, K. (2006). Keeping Transcriptional Activators under Control. Cell *127*, 261–264.

14. Li, H., Ban, F., Dalal, K., Leblanc, E., Frewin, K., Ma, D., Adomat, H., Rennie, P.S., and Cherkasov, A. (2014). Discovery of Small-Molecule Inhibitors Selectively Targeting the DNA-Binding Domain of the Human Androgen Receptor. J. Med. Chem. *57*, 6458–6467.

15. Li, T., Santockyte, R., Shen, R.-F., Tekle, E., Wang, G., Yang, D.C.H., and Chock, P.B. (2006). Expression of SUMO-2/3 Induced Senescence through p53- and pRB-mediated Pathways. J. Biol. Chem. *281*, 36221–36227.

16. Lim, M., Otto-Duessel, M., He, M., Su, L., Nguyen, D., Chin, E., Alliston, T., and Jones, J.O. (2014). Ligand-Independent and Tissue-Selective Androgen Receptor Inhibition by Pyrvinium. ACS Chem. Biol. *9*, 692–702.

17. Moore, R.D., and Morrill, G.A. (1976). A possible mechanism for concentrating sodium and potassium in the cell nucleus. Biophys. J. *16*, 527–533.

18. Ottis, P., and Crews, C.M. (2017). Proteolysis-Targeting Chimeras: Induced Protein Degradation as a Therapeutic Strategy. ACS Chem. Biol. *12*, 892–898.

19. Rytinki, M., Kaikkonen, S., Sutinen, P., Paakinaho, V., Rahkama, V., and Palvimo, J.J. (2012). Dynamic SUMOylation Is Linked to the Activity Cycles of Androgen Receptor in the Cell Nucleus. Mol. Cell. Biol. *32*, 4195–4205.

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

21. Sarge, K.D., and Park-Sarge, O.-K. (2011). SUMO and Its Role in Human Diseases. Int. Rev. Cell Mol. Biol. *288*, 167–183.

22. Small, E.J., Schellhammer, P.F., Higano, C.S., Redfern, C.H., Nemunaitis, J.J., Valone, F.H., Verjee, S.S., Jones, L.A., and Hershberg, R.M. (2006). Placebo-Controlled Phase III Trial of Immunologic Therapy with Sipuleucel-T (APC8015) in Patients with Metastatic, Asymptomatic Hormone Refractory Prostate Cancer. J. Clin. Oncol. *24*, 3089–3094.

23. Terry, C.A., Fernández, M.-J., Gude, L., Lorente, A., and Grant, K.B. (2011). Physiologically Relevant Concentrations of NaCl and KCl Increase DNA Photocleavage by an N-Substituted 9-Aminomethylanthracene Dye. Biochemistry (Mosc.) *50*, 10375–10389.

24. Tomlins, S.A., Laxman, B., Varambally, S., Cao, X., Yu, J., Helgeson, B.E., Cao, Q., Prensner, J.R., Rubin, M.A., Shah, R.B., et al. (2008). Role of the TMPRSS2-ERG Gene Fusion in Prostate Cancer. Neoplasia N. Y. N *10*, 177–188.

25. Wang, R.-T., Zhi, X.-Y., Zhang, Y., and Zhang, J. (2013). Inhibition of SENP1 induces radiosensitization in lung cancer cells. Exp. Ther. Med. *6*, 1054–1058.

26. Warburg, O. (1956). On the origin of cancer cells. Science *123*, 309–314.

27. Yang, S.-H., and Sharrocks, A.D. (2004). SUMO Promotes HDAC-Mediated Transcriptional Repression. Mol. Cell *13*, 611–617.

Appendix

Table A1.1. Dose-response experiment results for nucleotide analogs. Hill coefficients from IC_{50} analysis determination also presented.

	Name	Log IC_{50} (M)	Hill coefficient
1	ATP	-3.15	-0.93
2	2',3'-ddATP	-3.93	-0.98
3	(L)-2'-dATP	-3.68	-0.83
4	Ara-ATP	-3.70	-0.91
5	3'-dATP	-3.69	-0.91
6	3'-O-Methyl-ATP	-4.16	-0.96
7	2'-O-Methyl-ATP	-3.74	-0.92
8	2'-dATP	-3.69	-0.95
10	2'-amino-2'-dATP	-3.52	-0.92
11	2'-azido-dATP	-3.80	-0.95
12	2'-fluoro-dATP	-3.88	-0.93
13	3'-amino-ddATP	-3.12	-1.07
14	3'-azido-ddATP	-3.78	-0.92
15	2'-Bromo-dATP	-3.90	-0.92
16	2'-Iodo-dATP	-3.76	-0.85
18	8-chloro-2'-dATP	-2.47	-1.05
19	8-oxo-2'-dATP	-4.01	-0.65
20	N1-methyl-ATP	-2.85	-1.09
21	N6-methyl-2-amino-ATP	-3.61	-0.99
22	N6-methyl-ATP	-3.25	-0.96
23	2-amino-ATP	-3.56	-1.01
24	7-deaza-ATP	-3.38	-0.93
25	8-aza-ATP	-3.52	-0.95
26	8-azido-ATP	-3.76	-1.01
27	2-aminopurine-TP	-3.51	-0.95
28	2-FluoroATP	-3.33	-0.91
29	8-Bromo-ATP	-4.33	-0.91
30	8-Iodo-ATP	-3.73	-1.12
31	7-deaza-7-iodo-dATP	-3.99	-0.99
32	2-Methyl-thio ATP	-3.56	-0.88
33	6-thio-2'-dGTP	-4.34	-0.93

34	8-oxo-GTP	-3.15	-0.90
35	7-deaza-GTP	-3.36	-0.95
36	N2-methyl-2'-dGTP	-3.64	-0.98
37	N1-methyl-GTP	-3.83	-0.93
38	O6-methyl-GTP	-3.86	-0.93
39	GTP	-3.09	-0.91
40	dGTP	-3.72	-0.94
42	Ganciclovir-TP	-3.67	-1.06
43	2',3'-ddITP	-3.93	-0.98
44	ITP	-3.50	-0.96
45	2-amino-6-Chloro-Purine-TP	-3.92	-0.88
46	xanthosine-TP	-4.08	-0.86
47	6-chloro-purine-TP	-3.46	-0.89
48	thieno-GTP	-3.90	-0.97
49	α1-Thio-ATP	-3.91	-1.00
50	α1-Borano-2'-dATP	-3.73	-1.01
51	2-iodo-ATPgammaS	-4.49	-0.87
52	8-iodo-AppNHp ATP	-3.28	-1.20
53	2'-iodo-2'-deoxyadensine-5'-[(β,γ)-imido]triphosphate	-3.25	-0.96
54	GTP	-3.20	-1.01
55	GTPγS	-3.71	-0.94
56	ATP	-3.19	-0.90
58	5-nitro-1-indolyl-2'-dATP	-4.29	-1.02
59	2'-dZebularine-TP	-4.23	-0.90
60	Puromycin-triphosphate	-3.65	-1.08