Structural and functional studies on the regulation of the USP28 de-ubiquitinase and the SENP5 de-SUMOylase

Yang Zhen

Supervisor: Dr. David Reverter

Universitat Autonoma de Barcelona

Institut de Biotecnologia i Biomedicina (IBB)

September 2014

Universitat Autònoma de Barcelona

Institut de Biotecnologia i Biomedicina (IBB)

A thesis submitted in conformity with the requirements for the degree of Doctor in Institut de Biotecnologia i Biomedicina Universitat Autònoma de Barcelona

Semtember 2014

Thesis Director

Dr. David Reverter

PhD Candidate

Zhen Yang

Dedicated to

My family, especially my grandmother,

my friends and lab mates

Acknowledgements

First and foremost, I would like to appreciate my supervisor **Dr. David Reverter**, for his best training and patience to me; for the knowledge and scientific thinking he transferred to me; also for his dedication and guide on my project. **He promotes me to understand the secret and beauty of science!**

I would like to thank my best lab mates, Pablo, Kamela, Sergi, and Natalia, for their carefulness and endless support on my lab work. I am thankful to our neighbor group members: Susana, Ricardo, Virginia, Alba and Ana for their scientific advice and experience shared to me. I would also like to thank Bing Liu and Jara who directly helped me with my experiment.

I am also thankful to the members of other groups in our institute (IBB), for discussions, suggestions, and also happy atmosphere in the lab: Laia (one of my best friend in IBB), Sebastian (who is very nice and patient to help me to do the mammalian expression), Silvia (who help me to do the purification of protein) and our beautiful secretary Maite (who always help me to deal with the documents).

I want to sincerely thank Dr. Travis Stracker from IRB for his support on the mass spectrometry analysis and the advices of my paperwork. Thanks to Chinese Scholarship Council (CSC) for the grant.

I would like to thank my family members, my parents, my uncles, my aunts and **especially my grandmother** for their endless support and love. Finally, my love and thanks go to my dearest girlfriend Shan Huang, for her dedication, love and support in these years. I dedicate this thesis to them.

Zhen Yang

Barcelona, Autumn, 2014

Index

Introduction

1. General Introduction of Ubiquitin and SUMO

Post-translational modifications are one of the most effective ways by which evolution has increased the versatility in protein function using a relative paucity of genes. These covalent modifications, such as phosphorylation, acetylation or ubiquitination, rely on a set of enzymes for reversible conjugation/deconjugation that respond promptly to the requirements of the cell state and hence are essential for the regulation of cellular processes in a dynamic manner (Hunter *et al*, 2007). The covalent conjugation of Ub (ubiquitin) and Ubls (Ub like molecules), the best known of them being SUMO, has been intensely studied in the last few years. Ub and SUMO share a characteristic structural β-GRASP (β-Golgi reassembly stacking protein) fold (named the ubiquitin fold) and a similar conjugation mechanism, through their C-terminal carboxy group, to a particular lysine residue of their substrates. Despite these common structural and mechanistic traits, there are still some significant differences that allow them to be functionally classified into two separate groups (Denuc *et al*, 2010).

The post-translational modification of protein substrates by Ub (ubiquitin) and SUMO (small Ub-related modifier) contributes to signal transduction in pathways that control diverse events such as the cell cycle, apoptosis, cytokinesis, and protein degradation (Kerscher *et al*, 2006; Geiss-Freidlander and Melchior, 2007). Both Ubiquitylation and SUMOylation are conversable and regulated by the DUBs (Deubiquitylating enzymes) and SENPs (Sentrin/SUMO-specific proteases) in human. This highly reversibility of protein modifications enables the participation of proteins

regulated by them in multiple rounds of functional circuits and are also important to rapidly regulate and orchestrate protein functions in response to changes in a cell‘s state or its environment, without altering their synthesis or turnover rates.

The pervading influence of ubiquitin and the small ubiquitin related modifier SUMO on cellular pathways that control genome stability and the DNA damage response is being realized more and more and has become a focus of intense research over the past few years. Ubiquitin is perhaps best known for its role in targeting its substrates for degradation by the proteasome, and proteolytic functions of ubiquitylation are among those relevant for some DNA repair systems. However, the functional spectrum of both ubiquitin and SUMO is much wider, and many examples of non-proteolytic action of ubiquitin can be found among the various pathways dealing with DNA damage and genome maintenance. In cases where the consequences of the modifications have been elucidated in molecular detail, ubiquitin and SUMO mostly act as modulators of protein–protein or protein–DNA interactions, usually by providing or blocking binding sites for downstream effectors (Ulrich, 2009).

Given the pervasive influence of both ubiquitin and SUMO on cellular metabolism, it is not surprising to find situations where the two modification systems communicate. In fact, a growing number of proteins have been reported to serve as substrates for both ubiquitylation and SUMOylation, sometimes even involving the same lysine residue. As each modifier generally directs the target protein to a distinct fate, SUMO and ubiquitin are now often viewed as antagonists that control the properties of their substrates in a competitive manner (Ulrich, 2005).

2. Introduction of ubiquitin

2.1 Introduction of ubiquitylation

Ubiquitin is a small protein that is extremely well conserved among the eukarya but is absent from eubacteria and archaea. It is the most well-known molecular tag for proteolysis via proteasome (Ciechanover *et al*, 1998). It was discovered in 1975 as an 8.5kDa protein of unknown function (Schlesinger *et al*, 1975). In the following years, many studies on the Ub have been done and in 2004 the Nobel Prize was awarded to Aaron Ciechanover Hershko and Irwin Rose for their role in the ~~"~~discovery of ubiquitin-mediated protein degradation" (www.nobel prize.org).

Ub is activated and attached to substrate proteins by a series of enzymes. The trio of E1 ubiquitin-activating, E2 ubiquitin-conjugating and E3 ubiquitin-ligase enzymes perform an impressive array of ubiquitin modification reactions, including assembly of polyubiquitin chains. All eukaryotic species express multiple E2 and E3 isozymes, which can range up to several dozen E2s and many hundreds of E3s. This allows for the highly specific modification of many different proteins by ubiquitin, and such modifications are often under strict temporal and spatial control (Hochstrasser, 2009).

The C-terminal glycine of ubiquitin must be activated before it can form a covalent bond with another protein (Pickart *et al*, 2001; Huang *et al*, 2004) (Figure 1). Initially, the C-terminus is adenylated by E1, with the ubiquitin-AMP adduct remaining bound to the enzyme. An E1 cysteine side chain then attacks the ubiquitin C-terminus, yielding an E1-ubiquitin thioester intermediate. The activated ubiquitin is subsequently passed to the active site cysteine of an E2 (Lee *et al*, 2008). E2 proteins catalyze substrate ubiquitylation in conjunction with an E3 ligase. Ubiquitin E3s play

a paramount role in substrate recognition, although not all Ubl pathways necessarily require one. In the ubiquitin pathway, a different E3 may sometimes help add ubiquitins to a protein already modified by one or a few ubiquitins. Such E3s are sometimes called E4s, particularly when they are thought to extend a polyubiquitin chain.

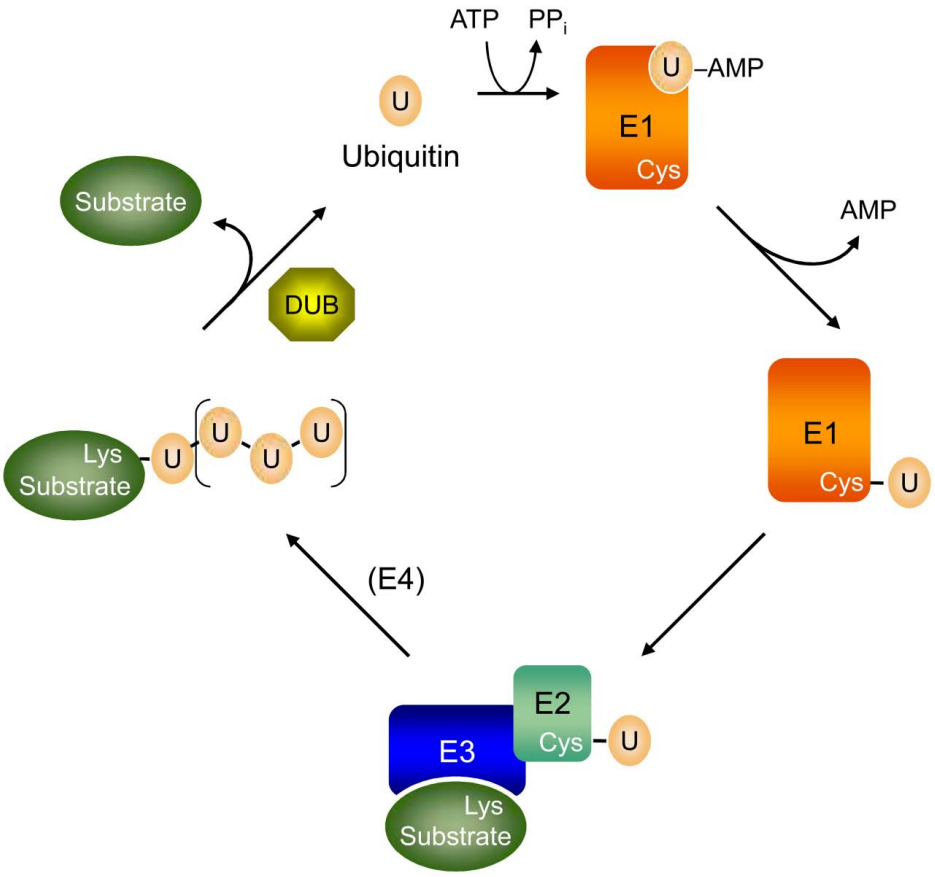


Figure 1 - The ubiquitin (U)-protein conjugation cycle. For ubiquitin (and at least some other Ubls), an E3 ligase is usually necessary to stimulate ubiquitin transfer from the E2 to a substrate, generally to a lysine ε-amino group. Additional ubiquitin molecules can be added either to other lysine side chains on the substrate or to ubiquitin itself, the latter leading to polymeric ubiquitin chains. Additional E3s can help assemble ubiquitin chains on substrates; when they act in this way, they are sometimes called ―E4s". Ubiquitin-substrate modifications are transient and can be removed by deubiquitylating enzymes or DUBs. In addition, ubiquitin and most Ubls are synthesized in precursorforms, and the C-terminal extensions are also removed by DUBs (Hochstrasser, 2009).

Ubiquitin is usually joined to proteins by an amide linkage between the C-terminus of ubiquitin and primary amino groups of the acceptor proteins (Pickart, 2001). The amine is most often a lysine ε-amino group, but it can also be the N-terminal N α-amino group (Ciechanover *et al*, 2004). In addition, recent work has shown *in vivo* ubiquitin attachment to cysteines, serines, and threonines in proteins (Cadwell *et al*, 2005; Ravid *et al*, 2007; Wang *et al*, 2007).

When ubiquitin forms polymers, the ubiquitin molecules are linked through the lysine side chain of one ubiquitin with the C-terminal carboxyl of the next ubiquitin. Ubiquitin has seven lysines, and all can contribute to such linkages, which also help dictate the fate of the modified substrate.

Depending on the linkage between ubiquitin molecules, these chains can encode distinct information. For example, chains linked through Lys48 of ubiquitin (K48-linked chains) are a targeting device for protein degradation by the 26S proteasome (Chau , *et al*, 1989; Thrower, *et al*, 2000), whereas K63-linked chains act as molecular scaffolds, bringing together subunits of oligomeric kinase or DNA repair complexes(Deng, *et al*, 2000; VanDemark, *et al*, 2001) (Figure 2).

As K48- and K63-linked ubiquitin chains were discovered many years ago, much has been learned about their functions, and they are often referred to as ―canonical" ubiquitin chains.

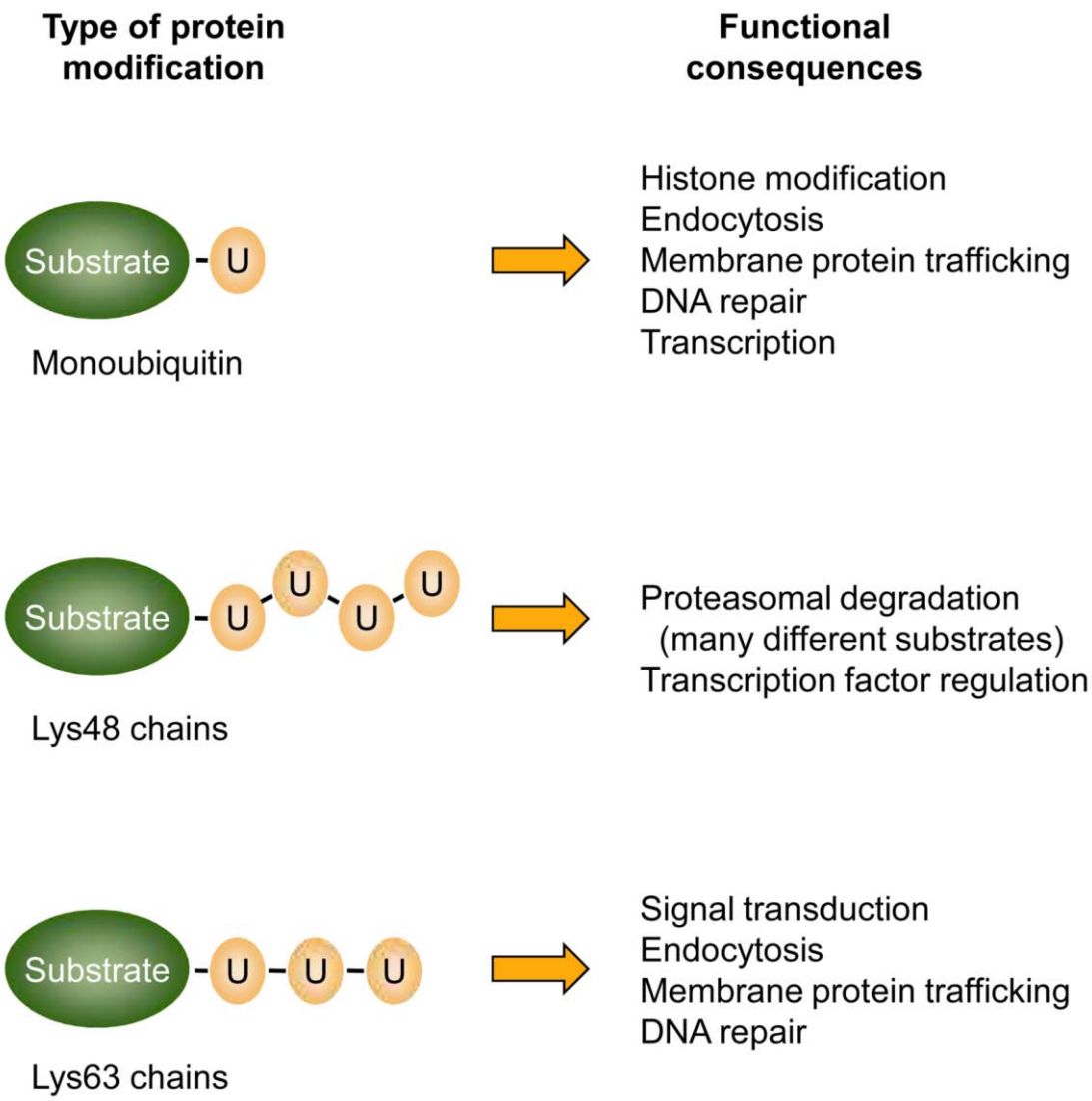


Figure 2 - Cellular processes that depend on ubiquitin conjugation. Protein attachment to a single ubiquitin allows recognition by a subset of ubiquitin-binding domains (UBDs) in target proteins, and this is important in the indicated general processes. Often a single, specific lysine is modified. Different polyubiquitin chains are thought to have different structures that allow discrimination among other UBDs, although other contextual cues, such as the cellular location where the modification occurs, may also help dictate the physiological consequences of the polyubiquitin attachment. Lys48-linked chains are most commonly associated with proteasomal binding and degradation. Not shown here are ubiquitin chains with mixed linkages or multi-site ubiquitylation of the substrate (Hochstrasser, 2009).

The Ub enzymes first catalyze the formation of an isopeptide bond between the C terminus of ubiquitin and usually a substrate lysine, leading to monoubiquitylation (Figure 3a). Monoubiquitylation can occur at a defined residue, such as Lys164 in proliferating cell nuclear antigen (PCNA) (Hoege, *et al*, 2002), or it might be confined to a domain, as in the transcription factor p53 (Carter, *et al*, 2007). It is possible that multiple lysine residues become modified with one ubiquitin each during multimonoubiquitylation (Figure 3b), with the epidermal growth factor receptor (EGFR) as an example (Haglund, *et al*, 2003).

Modification of the N terminus or one of the seven lysine residues of a substrate-attached ubiquitin leads to formation of polymeric chains. These chains can be short and contain only two ubiquitin molecules or long and incorporate more than ten moieties. Ubiquitin chains are homogenous if the same residue is modified during elongation, as in Met1- (or linear), Lys11-, Lys48-, or Lys63-linked chains (Figure 3c). Chains have mixed topology if different linkages alternate at succeeding positions of the chain (Figure 3d), as seen in NF-κB signaling or protein trafficking (Dynek *et al*, 2010; Gerlach *et al*, 2011; Boname *et al*, 2010; Goto *et al*, 2010). If a single ubiquitin is modified with multiple molecules, branched chains of unknown function are generated (Figure 3e).

All possible linkages have been detected in cells (Peng *et al*, 2003; Xu *et al*, 2009). For chains linked through Lys6, Lys27, Lys29, or Lys33, few substrates are known, and their significance is poorly understood. However, it has been well established that monoubiquitylation and four homogenous chain types trigger distinct

outcomes in the cell, suggesting that ubiquitylation can act as a code to store and transmit information.

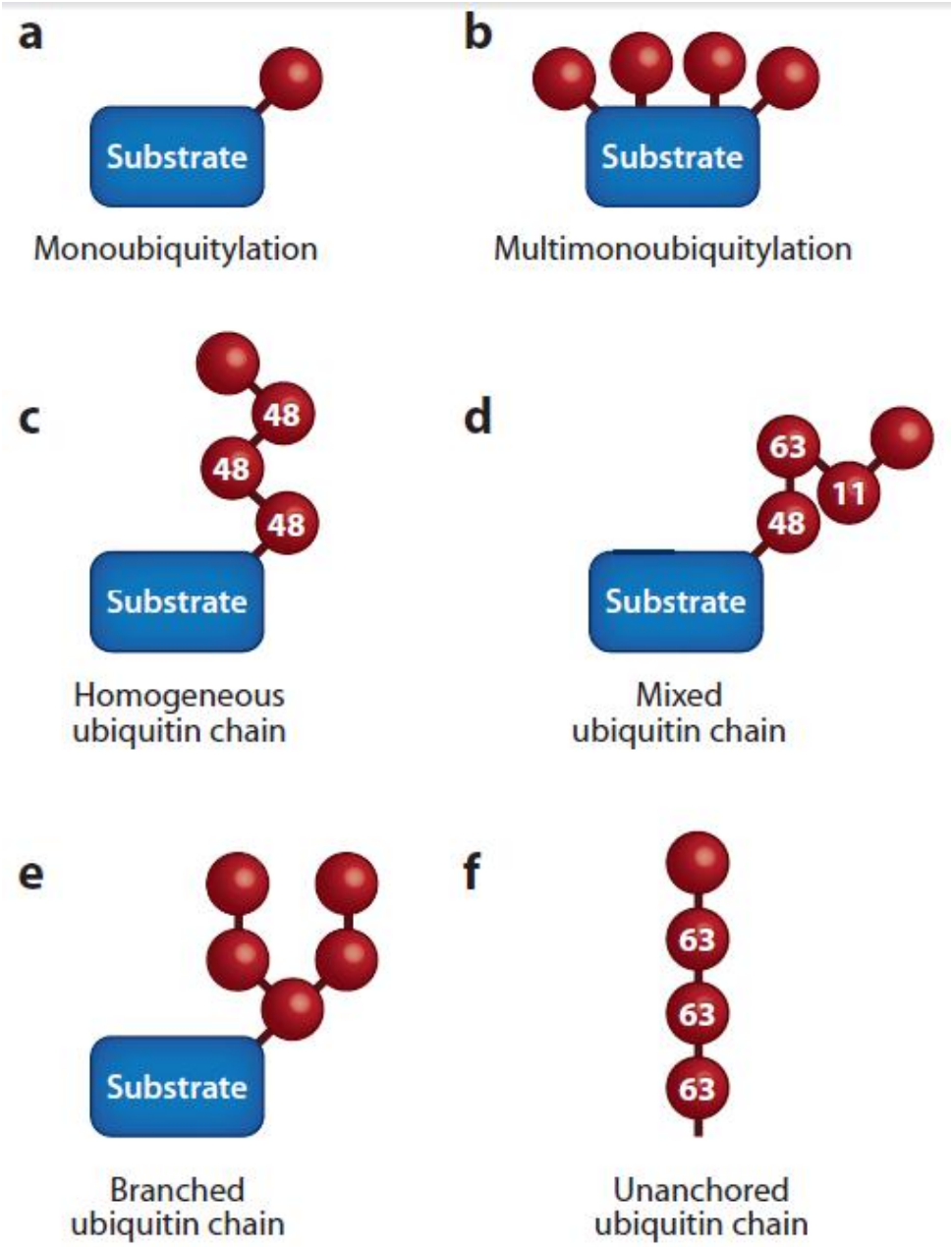


Figure 3 - The different topologies of ubiquitylation.

(a) Monoubiquitylation. (b) Multimonoubiquitylation. (c) Homogenous ubiquitin chain. (d) Mixed ubiquitin chain. (e) Branched ubiquitin chain. (f) Unanchored ubiquitin chain (Komander, *et al*, 2012).

We have to notice that a wide variety of proteins that contain a really interesting (RING) domain have been demonstrated to function as ubiquitin ligases that promote the transfer of ubiquitin from E2s to lysine residues in target proteins (Metzger, *et al*, 2014). Among the RING-type E3 ligases, RING finger protein 4 (RNF4 or SNURF) is proposed to ubiquitylated proteins that have been modified by SUMO and therefore is classified into a special category of E3s, termed as the SUMO-targeted ubiquitin ligases (STUbLs) (Perry, *et al*, 2008; Sriramachandran, *et al*, 2014).

Due to the activity of DUBs (Deubiquitylating enzymes), ubiquitin-modified proteins are only transiently modified (Amerik *et al*, 2004; Nijman *et al*, 2005). Dynamic modification of proteins by ubiquitin and other Ubls creates reversible switches between different functional states.

2.2 DUBs and De-ubiquitylation

Ubiquitylation plays a central role in degradation of proteins both through proteasomal targeting and by direct sorting to the lysosome. However, it is now becoming clear that reversible ubiquitylation is also a crucial mediator within intracellular signaling cascades as exemplified by Nuclear Factor- kB (NF-kB) signaling. Protein networks can be formed by interaction with specific ubiquitin binding domains of which there are at least 20 classes within the human genome (Dikic *et al*, 2009).

Reversibility of ubiquitylation is accomplished through DUBs. Approximately one hundred human DUBs fall into five classes (Figure 4): ubiquitin specific

proteases (USP), ubiquitin C-terminal hydrolases (UCH), ovarian tumour proteases (OTU), Josephins and the Jab1/MPN/MOV34 metalloenzymes (JAMM, also known as MPN1). The largest of the five families is the USP family (~55 members). In the case of the SUMO family of ubiquitin-like proteins, however, adducts are reversed by a specialized group of DUBs termed SENPs, all of which are cysteine peptidases (Komander *et al*, 2009; Reyes-Turcu, *et al*, 2009).The structure and function of these enzymes has recently been reviewed elsewhere (Hay *et al*, 2007 & 2010).

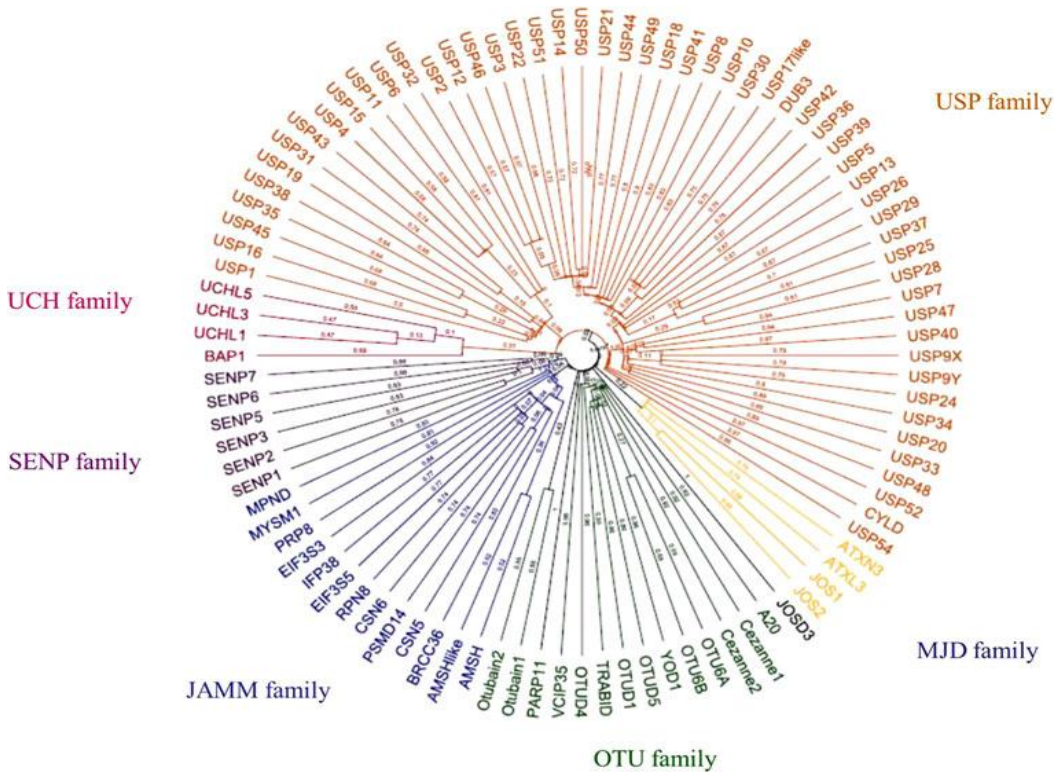


Figure 4 - DUB phylogenetic tree.

Approximately 100 genes belong to the DUB family of peptidases. Six classes of DUBs have been identified so far in the human genome. Five families belong to the cysteine peptidase class: the ubiquitin carboxylterminal hydrolases (UCH); the ubiquitin-specific proteases (USPs); the SENPs or SUMO peptidases; the OTU and the MJD. In addition, the MPN+/ JAMM family belongs to the metallo-peptidases class of enzymes. The phylogenetic tree represents only human DUBs and does not cover bacterial or viral DUBs that display additional levels of divergence (Jacq *et al*, 2013).

The crystal structures of a number of DUBs in the USP class have been resolved, including USP7/HAUSP, USP14, USP2, USP21 and USP8 providing the basis for molecular recognition studies of these proteases in the apo form and in complex with ubiquitin (Hu *et al*, 2002; Hu *et al*, 2005; Renatus *et al*, 2006; Avvakumov *et al*, 2006; Komander *et al*, 2008; Ye *et al*, 2011). These structural studies demonstrated that the mechanism for ubiquitin recognition is similar, which are homologous only within their catalytic site regions, and it was hypothesized that this recognition mechanism is common to all DUBs of the USP/UBP class (which is so far the most well characterized).

Genomics has identified at least 530 human genes that putatively encode enzymes involved in the conjugation and deconjugation of ubiquitin. Of these, at least 79 are thought to encode functional DUBs, some of which have multiple isoforms (Nijman *et al*, 2005; Wong *et al*, 2003). Considerable progress has been made in the study of ubiquitin conjugation, however, the study of DUBs, is still in its primary stages. Early research has been promising, implicating a number of DUBs, such as USP4 (UNP), USP6 (Tre-2), USP8 (UBPY), USP28 and UCHL5 (UCH37) in neoplastic disease (Jacq *et al*, 2013; Popov *et al*, 2007; Shah *et al*, 2009; Flügel *et al*, 2012; Wu *et al*, 2013; Diefenbacher *et al*, 2014).

The classical role of ubiquitin is to serve as a tag for protein destruction (Ciechanover, 1998). It follows that deubiquitylation can promote protein stability (Figure 5) and a diverse array of DUBs ensure some selectivity to this process.

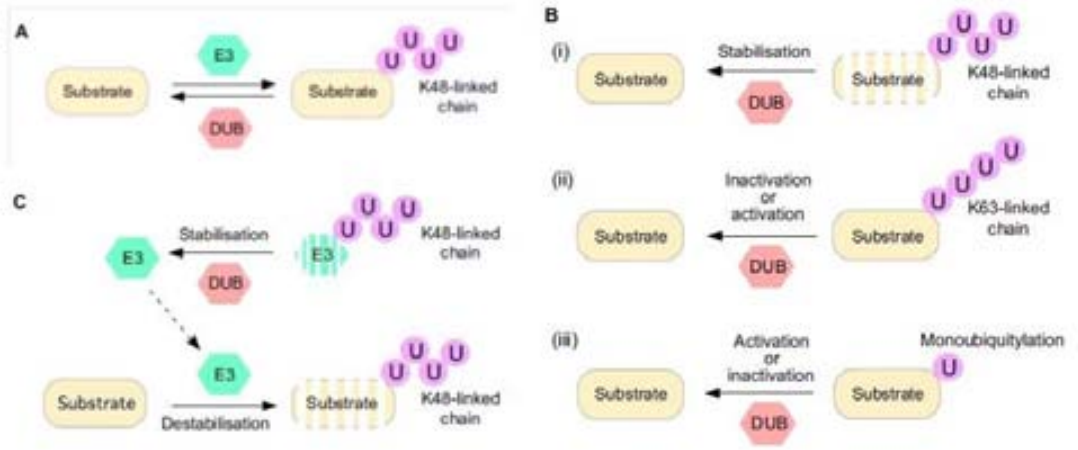


Figure 5 - DUBs regulate the stability or activity of proteins. (A) DUBs oppose the action of E3 ubiquitin ligases. (B) Depending on the type of ubiquitin chain that is attached to the substrate, DUBs can stabilise, or inactivate or activate their target protein. Examples from the text include the following DUB (substrate) configurations: (i) AMSH (EGFR), USP34 (AXIN), (ii) CYLD (DVL), (iii) USP9X (SMAD4). (C) Many E3 ligases undergo autoubiquitylation and can be stabilised by DUBs. In this scenario, DUBs may indirectly destabilise the protein that is targeted by the E3 ligase (Urbé *et al*, 2012).

In general terms, DUBs may influence the stability of key oncogenes or they may negatively regulate ubiquitin mediated signaling (Figure 6). Both oncogenic and tumour suppressive functions have been ascribed to individual DUBs. However, DUBs may have multiple substrates, thereby making it difficult to determine if a DUB has a net oncogenic or tumour suppressive function *in vivo* (Figure 6). Although knockout and overexpression models are helpful, there is evidence to suggest that function may vary between tissue types and stage of malignancy. For a subset of DUBs, mutations and/or altered expression in cancer specimens and cell lines have been described. Initiatives such as the Catalogue of Somatic Mutations in Cancer (COSMIC) (Forbes *et al*, 2008) and collation of gene expression data through

Oncomine (Rhodes *et al*, 2007) provide important resources for the association of DUBs with cancer.

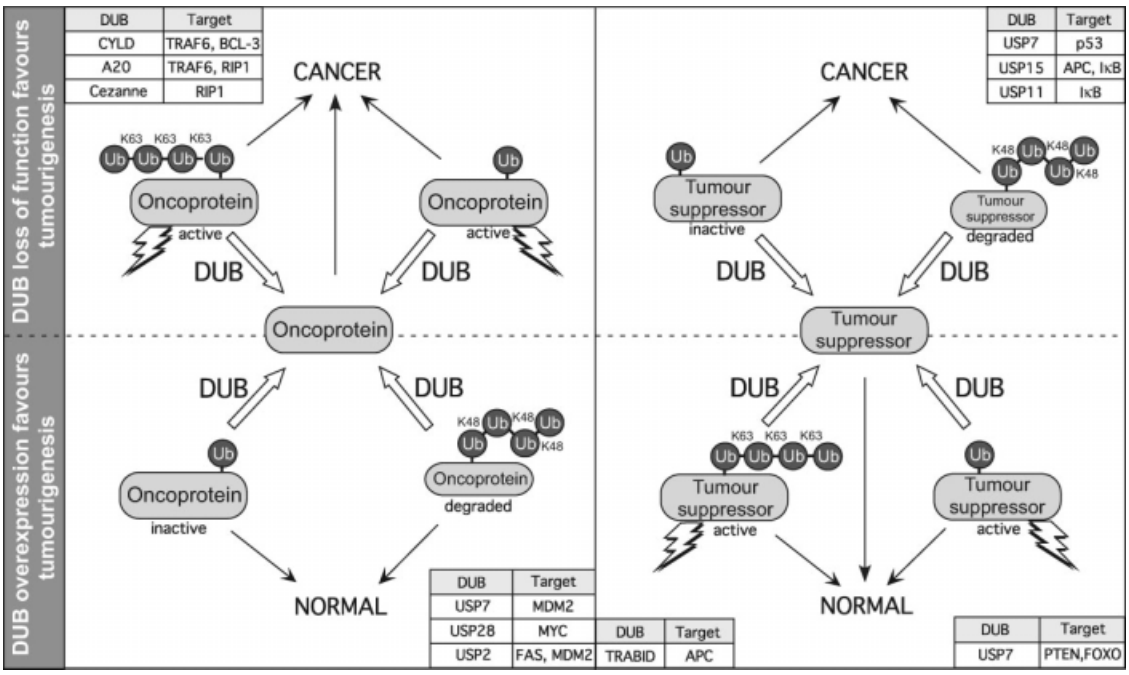


Figure 6 - Deubiquitinases are important regulators of oncogenes and tumour suppressors. Both overexpression and loss of function of DUBs can promote cancer. Ubiquitination of oncoproteins and tumour suppressors can promote their destabilization by targeting them for degradation (e.g., K48-linked poly-ubiquitination specifies proteasomal degradation), or regulate their activity (activation or inactivation). Activation here may refer to a variety of processes like translocation to the nucleus (e.g., PTEN and FOXO), or engagement in signalling protein interaction networks (TRAF6, RIP1). Specific DUBs implicated in tumourigenesis are shown with their cognate targets (Urbé *et al*, 2010).

2.3 De-ubiquitylation of the Ub chains

To specifically control ubiquitin-dependent signaling, the DUBs have to deal with chains of distinct linkage, topology, and length.

Several DUBs, referred to as housekeeping enzymes, play important roles in establishing the ubiquitin code. For example, proteasome-bound DUBs, such as

USP14, UCH37/UCHL5, and RPN11/POH1, protect ubiquitin from degradation (Finley, 2009).This process is vital for keeping sufficient levels of free ubiquitin that can be used for chain assembly. Similar functions might be performed by DUBs that interact with ubiquitin-processing complexes, such as the COP9 signalosome (USP15) (Hetfeld, *et al*, 2005), or the p97 segregase [YOD1 (Ernst, *et al*, 2009), VCIP135 (Uchiyama, *et al*, 2002), Ataxin-3 (Madura, *et al*, 2003)].

Another large group of DUBs disassembles chains independently of the linkage, yet these enzymes gain specificity by being targeted to a select set of substrates. These DUBs include most members of the USP family, which regulate many cellular reactions, including splicing, protein trafficking, or chromatin remodeling.

Many USP DUBs are recruited to substrates through interaction domains (Komander, *et al*, 2009) or adaptor subunits (Sowa, *et al*, 2009). Although a comprehensive analysis has not been reported, most USPs are active against all linkages (Ye, *et al*, 2011; Virdee, *et al*, 2010; Barford, *et al*, 2011) and also hydrolyze the isopeptide bond between the substrate and the first ubiquitin. An exception from this nonspecificity is CYLD, which prefers Met1- and Lys63-linked chains (Barford, *et al*, 2011; Xia, *et al*, 2009; Hofmann, *et al*, 2008). Hence, most USPs can be considered nonspecific with regard to the ubiquitin code but specific with respect to their substrates. In contrast to the aforementioned examples, several DUBs respond to the ubiquitin code and display specificity toward one or a few linkages. JAMM family DUBs are often Lys63 specific, as seen for AMSH (McCullough, *et al*, 2004), AMSHLP (Sato, *et al*, 2008), BRCC36, and POH1 AMSHLP (Cooper, *et al*, 2009). In

addition, linkage-specific OTU DUBs have been described; these descriptions showed that OTUB1 is specific for Lys48 linkages (Edelmann, *et al*, 2009; Wang, *et al*, 2009), Cezanne is specific for Lys11 linkages (Bremm, *et al*, 2010), and Trabid is specific for Lys29 and Lys33 linkages (Virdee, *et al*, 2010). As linkage-specific DUBs may not be able to cleave off the last ubiquitin (Bremm, *et al*, 2010), their activity might generate monoubiquitylated substrates with distinct signaling properties.

The structure of AMSH-LP with Lys63- linked diubiquitin revealed the basis for the linkage specificity displayed by JAMM family DUBs (Sato, *et al*, 2008). AMSH-LP binds to the open conformation of the Lys63-linked diubiquitin and contacts Gln62 and Glu64 of the proximal ubiquitin. Thus, reminiscent of some UBDs, JAMM DUBs might recognize the sequence context of the Lys63-isopeptide bond. The structure of Trabid revealed a different mechanism to achieve specificity; this enzyme uses an Ankyrin-repeat UBD directly upstream of the catalytic OTU domain to position the proximal ubiquitin (Licchesi, *et al*, 2011). However, for the OYU and remaining DUB families, structures are available only in complex with a single ubiquitin (Hu *et al*, 2002; Weeks, *et al*, 2011; Johnston, *et al*, 1999; Messick, *et al*, 2008), which revealed high-affinity binding sites for the distal ubiquitin, while leaving the interaction sites for the proximal ubiquitin that provides the modified lysine unclear. These structures did imply, however, that compact chain conformations should not be recognized by DUBs unless they undergo extensive remodeling to expose the isopeptide bond.

2.4 Introduction of USP25 and USP28

USP25 and USP28 are the members of a family of deubiquitinating enzymes (DUBs). It belongs to the USP family, which comprises more than 50 members (Komander, *et al*, 2009; Reyes-Turcu, *et al*, 2009). USP28 is highly homologous to USP25, which has been biochemically characterized. Searches in silico predicted that both USP28 and USP25 contain an Ubiquitin-Associated Domain (UBA) and two ubiquitin interaction motifs (UIM) in the N-terminal region of the proteins.

USP25 contains an ubiquitin-associated domain (UBA) and two ubiquitin interaction motifs (UIMs). Although several DUBs contain ubiquitin-binding domains, their role in DUB function is largely unknown. It is reported that USP25 is a target for SUMO conjugation that is preferentially conjugated with SUMO2/3 as a consequence of its favorable binding to SUMO2/3, compared to SUMO1. Two SUMOylation sites are identified within the first and directly adjacent to the second ubiquitin interaction motif. These UIMs are required for the full catalytic potential of USP25 toward ubiquitin chains. Importantly, SUMOylation of USP25 impairs its activity as a result of reduced affinity to ubiquitin chains (Meulmeester *et al*, 2008). On the other hand, another study showed that USP25 catalytic activity did not strictly depend on the UBDs, but required a coiled-coil stretch between amino acids 679 to 769. USP25 oligomerized but this interaction did not require either the UBDs or the C-terminus (Denuc *et al*, 2009).

USP28 was identified through its homology to USP25 (Valero *et al*, 2001) and subsequently found as an interaction partner of 53BP1, a key regulator of DNA repair

pathway choice (Zhang *et al*, 2006). The catalytic activity of USP28 was reported to be required for IR-induced apoptosis and the stability of numerous DDR regulators (Zhang *et al*, 2006). Independently, USP28 was reported to stabilize the Myc proto-oncogene by antagonizing the activity of the SCF^{FBW7} ubiquitin ligase complex (Popov *et al*, 2007). This function of USP28 was required for Myc induced transformation and it was found to be upregulated in human colon carcinoma and important to prevent differentiation. It was proposed that the dissociation of USP28 from Fbw7 in response to DNA damage provides a potential mechanism that couples Myc stability to DNA damage.

Subsequent work has confirmed the interaction between USP28 and 53BP1 but found only minor effects on the DDR and no impairment in 53BP1 dependent processes, suggesting that it may not represent an attractive therapeutic target for chemosensitization (Jacq *et al*, 2013 and Knobel *et al*, 2014). However, its conditional depletion in a mouse model of colorectal cancer led to a significant increase in tumor latency, suggesting that in particular contexts, the modulation of its activity is likely to be highly relevant to cancer (Diefenbacher *et al*, 2014).

Although SUMO is not a direct tag for proteosomal degradation, there are several examples for the crosstalk between the SUMO and the ubiquitin modification systems (Praefke *et al*, 2011). For example, in the case of IkBα (Inhibitor of transcription factor NF-κB) and PCNA (Proliferating Cell Nuclear Antigen) SUMO or ubiquitin are conjugated on particular lysine residues and thus determine the protein fate in the cell (Desterro *et al*, 1998; Hoege *et al*, 2002). In this context of particular

interest was the discovery of ubiquitin-dependent degradation by specific SUMO-target ubiquitin E3 ligases (STUbL) that can recognize substrates with polySUMO chains (Prudden *et al*, 2007; Sun *et al*, 2007; Xie *et al*, 2007; Uzunova *et al*, 2007). Another example includes the ubiquitin E2 conjugating enzyme E2-25K, in which SUMO conjugation prevents interaction with the ubiquitin E1 enzyme (Pichler *et al*, 2005). Finally, the DUB protease USP25 has been shown to be either SUMOylated or monoubiquitined on Lys99, and at least for the model substrate MyBPC1 (Myosine-Binding Protein C) ubiquitin modification enhances the deubiquitinating activity, whereas SUMO modification has an opposite effect (Denuc *et al*, 2009). Interestingly, the ubiquitin-binding domains at the N-terminal region of USP25 seem to play a role in the regulation of the protease activity (Meulmeester *et al*, 2008; Denuc *et al*, 2009).

3. Introduction of SUMO

3.1 Introduction of SUMOylation

The ubiquitin-related protein SUMO-1 was discovered in studies on nuclear import in mammalian cells as a covalent modification of RanGAP1. This discovery may have been facilitated by its unique property of being nearly quantitatively and constitutively modified with SUMO. This modification targets the otherwise cytosolic RanGAP1 to the nuclear pore complex (NPC) where it participates in nuclear import by activating the GTPase activity of the cytosol/nucleus shuttling factor Ran (Matunis *et al*, 1996; Mahajan *et al*, 1997; Mahajan *et al*, 1998; Lee *et al*, 1998; Matunis *et al*, 1998; Saitoh *et al*, 1998). SUMOylation of RanGAP1 leads to its interaction with the Ran binding protein RanBP2 at the cytoplasmic filaments of the NPC. At the same time, RanBP2 itself is also modified by SUMOylation and, moreover, has recently been shown to act as a SUMO ligase (Saitoh *et al*, 1998; Pichler *et al*, 2002). SUMO was independently identified in a variety of studies explaining why in the literature it also appears as dGMPT, dPIC1T, dsentrinT, dSMT3T, or dUBL1T(Matunis *et al*, 1996; Boddy *et al*, 1996; Okura *et al*, 1996; Meluh *et al*, 1995; Shen *et al*, 1996).

SUMO is encoded by single genes in yeast species and invertebrates. Whereas the SUMO-encoding SMT3 gene is essential in *Saccharomyces cerevisiae* (Johnson *et al*, 1997), its counterpart pmt3 in *Schizosaccharomyces pombe* is not (Tanaka *et al*, 1999). Pmt3_mutants, however, grow poorly and display various phenotypes that have also been associated with conditional mutations in the SUMO system of *S. cerevisiae*. Four different SUMO isoforms termed SUMO-1, SUMO-2, SUMO-3, and SUMO-4 have been detected in mammals. SUMO-2 and SUMO-3 are very similar in sequence

and therefore sometimes termed SUMO-2/3 in one breath. The divergence of the functions of these isoforms is just beginning to emerge. SUMO-1，which displays a slightly higher degree of similarity to yeast SUMO/Smt3 (~47% identical residues), seems to be the most prominently conjugated isoform under normal conditions. SUMO-2/3 (~45% of the residues identical to those of Smt3) appears to be preferentially conjugated to proteins under stress conditions such as increased temperature (Saitoh *et al*, 2000). There are, however, examples of substrates such as topoisomerase II and CAAT/enhancer-binding protein-beta (C/EBPh) that are specifically modified by SUMO-2/3 under normal physiological conditions (Eaton *et al*, 2003; Azuma *et al*, 2003). The recently identified fourth isoform, SUMO-4, is encoded by a sequence that lies within an intron of the human TAB2 gene (Bohren *et al*, 2004). The expression of this gene is in kidney cells. While SUMO-2, SUMO-3, and SUMO-4 share a SUMO attachment consensus site (see below), such a site is absent from SUMO-1. Consistent with this observation, in contrast to SUMO-1, SUMO-2/3 as well as SUMO-4 have been shown to form SUMO chains *in vitro* and *in vivo*(Bohren *et al*, 2004; Tatham *et al*, 2001). Multiple SUMO isoforms, eight to be exact, are encoded by the genome of the model plant *Arabidopsis thaliania*. Similar to mammals, the conjugation of certain isoforms (SUMO-1 and SUMO-2) is induced when *Arabidopsis* is subjected to heat stress (Bohren *et al*, 2004). The structure of human SUMO-1 has been determined by NMR and compared to the crystal structure of ubiquitin (Bayer *et al*, 1998). More recently, the structure of budding yeast SUMO (Smt3) has been determined after co-crystalization with Ulp1 (Mossessova *et al*,

2000), and in solution with and without Ubc9 (Sheng *et al*, 2002).

Although the sequence identity between SUMO and ubiquitin is relatively low (~18% identity) the overall three-dimensional structures are very similar (Figure 7).

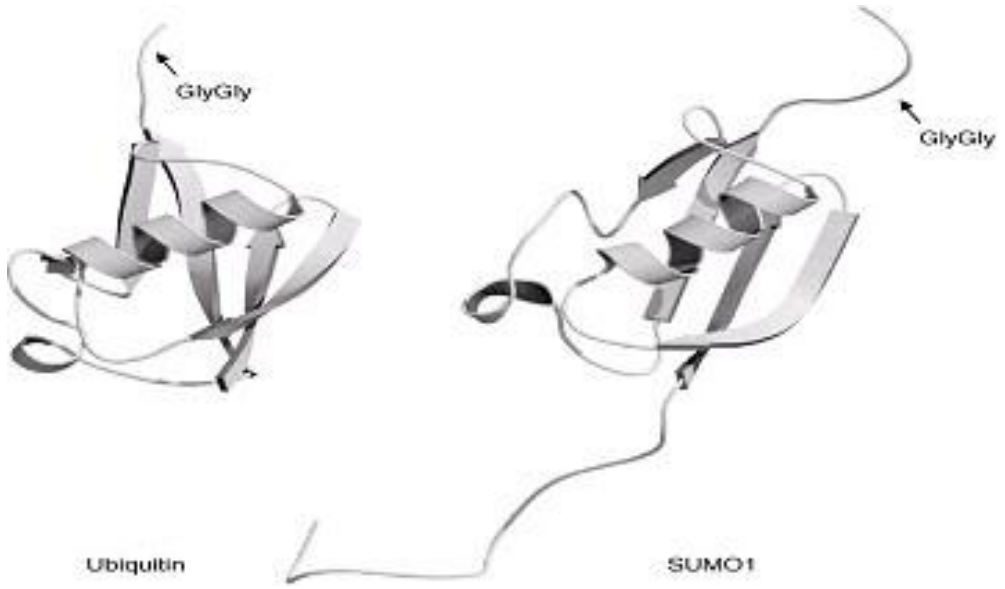


Fig 7 - The comparison between SUMO and UB

The surface charge distributions of the two proteins, however, are quite different, indicating that they interact specifically with distinct enzymes and substrates. Another prominent feature of SUMO-1 is a protruding long and flexible N-terminal domain, which is absent in ubiquitin. In yeast, a lysine residue within this N-terminal domain has been implicated in the formation of poly-SUMO chains (Bylebyl *et al*, 2003; Bencsath *et al*, 2002). Surprisingly, however, the entire extension including this lysine can be deleted without severe consequences for the yeast indicating that, in contrast to ubiquitin, chain formation is not important for SUMO function in *S. cerevisiae* (Bylebyl *et al*, 2003). A feature that is shared between the mature forms of SUMO and

ubiquitin, and also some other Ubls, is a di-glycine motif at the C-terminus. This motif was shown to be critical for SUMO conjugation in *S. cerevisiae* (Johnson *et al*, 1997).

3.2 The SUMOylation Pathway

There is a remarkable conservation of the mechanisms between ubiquitination and SUMOylation pathways. Like ubiquitin, SUMO can become covalently attached to a protein through an energy-consuming reaction cascade that requires the consecutive action of up to three enzymes: an E1 activating enzyme (the heterodimer Aos1–Uba2), an E2 conjugating enzyme (Ubc9) and one of several SUMO E3 ligases (Figure 8). SUMO reacts with the free amino group (NH2) of a lysine amino-acid residue on its target protein to form an isopeptide bond. Although SUMOylation usually results in the attachment of single SUMO moieties to one or a few acceptor lysine residues in the target protein, in some cases SUMO chains can form. This requires SUMO–SUMO isopeptide bond formation, which has been observed for SUMO in yeast and for SUMO2/3 in mammalian cells. Experiments in yeast suggest that SUMO chain formation is not essential for an organism‘s growth, but is required for spore formation in meiotic cell division.

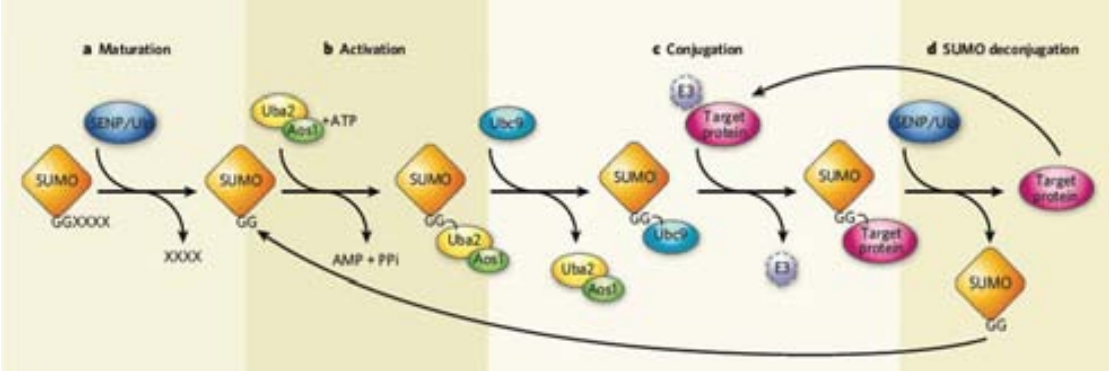


Figure 8 - SUMO conjugation and deconjugation. Newly synthesized SUMO is ‗immature' as it cannot conjugate to its targets until two glycine residues (GG) close to its carboxy terminus are exposed in a reaction that removes some carboxy-terminus residues and is mediated by SUMO-specific proteolytic enzymes — SENP proteases in mammals and Ulp proteases in yeast. Mature SUMO is then activated in an energy (ATP)-consuming step by the E1 activating enzyme Aos1–Uba2. Subsequently, SUMO is transferred to the E2 conjugating enzyme Ubc9. In the last step, the carboxyl group of the glycine residue at SUMO's carboxy terminus forms an isopeptide bond with the amino group of a lysine residue in its target protein. This step is usually facilitated by E3 ligases, but some targets are efficiently SUMOylated by E2 alone. SUMOylation is reversible, because the SUMO-specific proteases of the SENP/Ulp family also efficiently cleave the isopeptide bond between SUMO and its target. Both the released SUMO and the target protein then become available for subsequent rounds of modification. (Meulmeester, 2008)

The ubiquitin E1 exists as a monomer while SUMO E1 exists as a heterodimer, though both individual components are related to the ubiquitin enzyme; the Sae1 subunit (also called Aos1) of SUMO E1 resembles the N-terminal of ubiquitin E1 while the Sae2 subunit (also called Uab2), where the catalytic cysteine is located, resembles the C-terminal (Azuma *et al*, 2001). Structural analysis of E1 heterodimer indicated that SUMO interacts exclusively with subunit Sae2, and Ubl recognition may be dependent on conserved residues within this subunit (Tong *et al*, 1997). Though SUMO1 E1 enzyme exists as two distinct subunits, the individual

components are only found as part of the heterodimer within the cell.

The SUMO E2, Ubc9, also shares sequence similarity with its ubiquitin counterpart and the two enzymes share essentially the same folded structure (Hay *et al*, 2001). Ubc9 also shows a strong, unique overall electrostatic dipole which might contribute to its ability to recognize and conjugate SUMO to its substrates without the help of E3 ligases. SUMO E2 specifically recognizes substrates containing the motif ψKxE/D, where ψ is a large hydrophobic residue, K is the acceptor lysine and x is any amino acid (Bernier-Villamor *et al*, 2006). SUMO E2 makes direct contact with this motif and structural analyses of the E2 in complex with SUMO substrates have revealed a hydrophobic pocket on E2 that accommodates the acceptor lysine, with residues immediately flanking this lysine also facilitating substrate recognition via hydrogen bonding and electrostatic interactions with the surface of E2 (Hay *et al*, 2001; Bernier-Villamor *et al*, 2006; Sampson *et al*, 2001).

In general there are three types of E3 SUMO ligases: those belonging to the Siz/PIAS (protein inhibitor of activated STAT) family; those containing a domain in the large vertebrate nuclear protein RanBP2/Nup358; and the polycomb group protein Pc2 (Jackson *et al*, 2001; Hochstrasser *et al*, 2001; Shuai *et al*, 2000; Wotton *et al*, 2003; Bencsath *et al*, 2002). The members PIAS family of proteins all share a ~400 residue N-terminal domain that contains smaller regions of similarity but more importantly it contains a SP-RING domain thought to function in a manner analogous to the ubiquitin RING E3 ligases by providing a scaffold for conjugation by bringing Ubc9 and the substrate together. This family includes Siz1, Siz2 and methyl

methanesulphonate-sensivity protein 21 (Mms21) in yeast and PIASxα, PIASxβ and PIAS y in humans. The RanBP2 ligase consists of a ~300 residue domain called the IR (internal repeat) and can be found in the core protein RanBP2 (Bencsath *et al*, 2002). This domain contains two internal repeats of a ~50 residue sequence which not only function as SUMO ligases but are also responsible for localizing RanGAP1-SUMO to the nuclear pore (Matunis *et al*, 1998). While the RanBP2 and PIAS seem to have non-redundant cellular functions with some substrates are only SUMOylated exclusively by one ligase or the other there have been cases where SUMOylation can be induced by both (Miyauchi *et al*, 2002; Sobko *et al*, 2002; Kannouche *et al*, 2004). The polycomb group (PcG) protein Pc2 is the last group of reported SUMO E3 enzymes. In humans, these proteins form large multimeric complexes (PcG bodies), which are involved in the stable maintenance of transcriptional repression and Pc2 has been shown to induce SUMOylation of various types of proteins within this pathway (Schuettengruber *et al*, 2007; Ringose *et al*, 2007; Satijn *et al*, 1997; Jacobs *et al*, 2002). Moreover overexpression of Pc2 in cells causes SUMO and Ubc9 to localize to PcG bodies, suggesting that PcG bodies may be major sites of SUMOylation.

Besides the aforementioned SUMO consensus motifs, the SUMO conjugation machinery also makes use of SIMs or SUMO Interaction Motifs, to help recognize SUMO. SIMs are generally characterized by a hydrophobic core (V/I) X (V/I) (V/I), flanked at either its N or C terminus by acidic amino acids (Hecker, C *et al*, 2006). When in complex with SUMO, the SIM adopts a parallel or anti-parallel β-strand

conformation, which allows the hydrophobic side chains of the SIM to occupy a hydrophobic pocket on the SUMO surface (Reverter *et al*, 2005). SUMO E3s, such as those containing the Ran-binding protein 2 (RanBP2) domain and those belonging to the Siz/PIAS family of E3 ligases, are known to have both E3 ligase elements as well as SIMs (Rytinki *et al*, 2009; Pichler *et al*, 2004). Of the SIM-containing E3 ligases that have been characterized, some display substrate preferences, such as those mentioned above, while others show preference for specific SUMO-isoforms, such as in the case of ubiquitin-specific protease 25 (USP25) and transcriptional regulator COREST1, which both exhibit a preference for SUMO2/3ylation (Meulmeester *et al*, 2008; Ouyang *et al*, 2009).

3.3 Introduction of SENP protease family

SUMOylation is a dynamic process that is readily reversed by a family of ubiquitin-like protein-specific proteases (Ulp) in yeast and Sentrin/SUMO-specific proteases (SENP) in human.

During SUMO metabolism, Ulp/SENPs catalyze three distinct processes: processing, deconjugation, and chain editing (Drag *et al*, 2008). Like ubiquitin, SUMO proteins are expressed as precursor proteins that carry a C-terminal extension of variable length (2–11 amino acids) found after a conserved di-glycine motif. To function as a modifier of target proteins, the C-terminal di-glycine motif of the SUMO proteins must be exposed by the action of SUMO specific protease (Jentsch *et al*, 2000). SUMO processing activity of SENPs is responsible for cleavage after the C

terminal di-glycine motif and SUMO deconjugation activity of SENPs is required for the cleavage of amide bond between the C-terminus of the mature SUMO and the ε-amino group of the target lysine within the substrates. Chain editing refers to the dismantling of SUMO chains. (Figure 9)

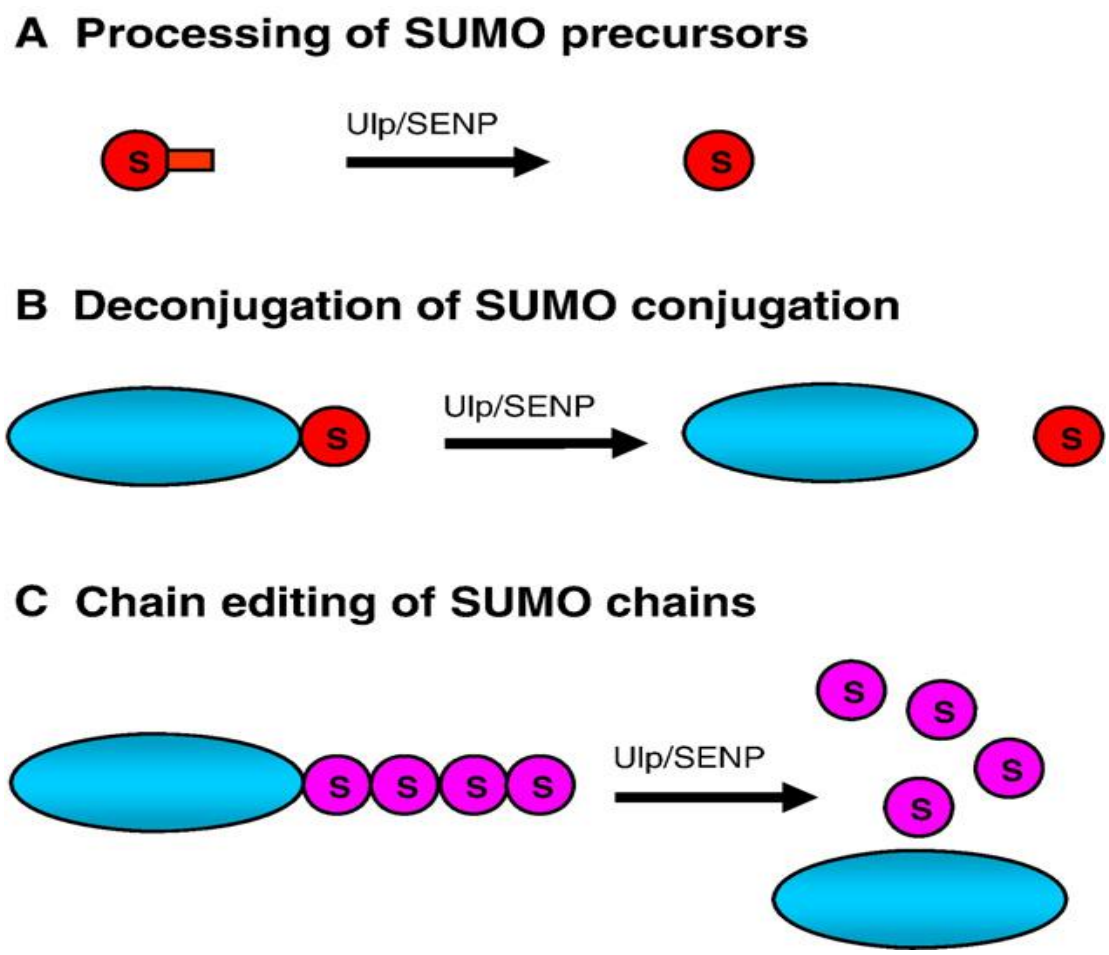


Figure 9 - Catalytic functions of deSUMOylating enzymes. (A) Processing of SUMO precursors by Ulp/SENPs generate free SUMO monomers that are ready for protein modification. (B) Reversal of SUMO conjugation occurs by deconjugation of SUMO from modified proteins. (C) Ulp/SENPs function in the editing of SUMO at the distal ends of poly-SUMO-2/3 chains conjugated to protein substrates. The red circle represents SUMOs and the pink circle represents SUMO-2/3. (Jung Hwa Kim *et al*, 2009)

SUMO-specific proteases are C48 cysteine proteases that possess a conserved catalytic domain characterized by the catalytic triad (histidine, aspartate, and cysteine) and a conserved glutamine residue required for the formation of the oxyanion hole in the active site (Li *et al*, 1999). Members of the C48 cysteine protease family have N- and C-terminal sequences that differ from each other. Homologs of these proteases are present in plant, yeast, and mammalian cells.

In *S. cerevisiae*, two Smt3-specific proteases, Ulp1 and Ulp2, have been characterized, and both enzymes can de conjugate Smt3 from modified proteins and process Smt3 precursors to the mature form with the C-terminal diglycine (Li *et al*, 1999; Takahashi *et al*, 2000). Comparison of Ulp1 and Ulp2 reveals that homology between the two enzymes is confined to a protease domain of _200 amino acids. SENPs are cysteine proteases that belong to the family of proteases typified by the adenovirus protease. Database searching initially identified seven genes for human proteins with significant sequence homology to yeast Ulp1; these genes were believed to encode SUMO (also known as Sentrin) proteases (i.e. SENPs) (Yeh 2000). Sequence comparisons of these genes with that of Ulp1 and Ulp2 (Figure 10a) indicates that SENP1 (known as SUMO protease 2, SuPr2, in mouse) (Gong *et al* 2000), SENP2 [also designated Axam, SuPr1, SUMO specific protease 3 (SSP3) or Smt3 specific isopeptidase 2 (SMT3IP2)], SENP3 (SMT3IP1 in mouse) and SENP5 (Zhang *et al* 2002; Nishida *et al* 2000; Hang and Dasso 2002; Girdwood *et al* 2003; Kadoya *et al* 2002; Best *et al* 2002; Gong 2006; Di Bacco *et al* 2006) are related to Ulp1, whereas SENP6 (also designated SUMO specific protease 1, SUSP1) (Kim *et al*

2000; Mukhopadhyay *et al* 2006) and SENP7 are related to Ulp2. Recently, an additional SENP, SUSP4, was detected in mouse cells. Although it does not seem to have a direct human counterpart, it is most closely related to SENP2 (Lee *et al* 2006). NEDP1 (NEDD8 protease 1, also known as SENP8) was found to be on a distinct evolutionary branch, consistent with its role as aNEDD8-, rather than SUMO-, specific protease (Mendoza *et al* 2003; Gan-Erdene *et al* 2003; Wu *et al* 2003). Although the SENPs all share a similar C-terminal catalytic domain, the N-terminal regions are largely unrelated (Figure 10b) but seem to be involved in directing these proteins to distinct subcellular localizations (Melchior *et al* 2003).

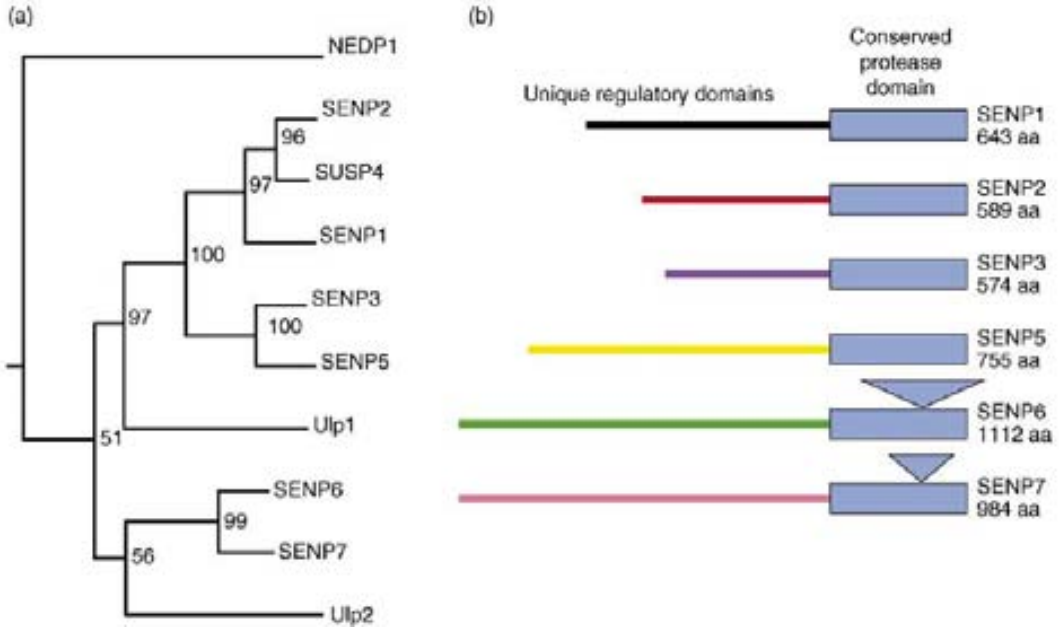


Figure 10 - The family of SENPs. (a) Phylogenetic relationship between members of the SENP family from yeast and humans. A distance tree shows the relationship between SENPs from budding yeast (Ulp1 and Ulp2) and humans (SENP1, SENP2, SENP3, SENP5, SENP6, SENP7 and SUSP4), including the human form of the NEDD8-specific protease (NEDP1), which is structurally related to the SENPs but does not cleave SUMO. To generate the tree, the PHYLogeny Inference Package (PHYLIP) software package was used (Felsenstein, J. 1995), and displays bootstrap values (as a percentage), essentially indicating confidence in the existence of that ancestral node. The tree is midpoint rooted, which puts NEDP1 as the outgroup on a branch between the two yeast SENPs. Assuming divergent evolution, this suggests that the NEDD8 protease diverged from the SENPs

relatively early in eukaryotic evolution. (b) Domain organization of SENPs. Both SENP6 and SENP7 contain inserts in the catalytic domain. Alignment of this sequence with the known structure of SENP1 and SENP2 suggests that these inserts will not disrupt the catalytic domain and will represent an additional domain that projects from the surface of the protein (Hay 2007).

3.4 Introduction of SENP3 and SENP5

SENP1 localizes to the nucleus, excluding the nucleolus (Gong *et al* 2000), and that SENP2, also called SuPr1, localizes to the nucleoplasmic face of the nuclear pore complex (Zhang *et al* 2002； Hang *et al* 2002). Both SENP1 and SENP2 are able to remove SUMO from all SUMO-modified proteins. The structures of a catalytically inert form of the human SENP2 catalytic domain in complex either with human SUMO-2 and SUMO-3 precursors or with conjugated human RanGAP1 containing either SUMO-1 or SUMO-2 were determined. These structures suggest how SUMO processing and deconjugation substrates interact with the protease before cleavage, and they identify additional surfaces on the protease that are important in SENP2-mediated reactions (Reverter *et al* 2006).

SENP3 and SENP5 constitute a subfamily of SENPs that share considerable sequence homology and exhibit similar substrate specificities. Both of them are localized at the nucleolus and they also have specificity for SUMO-2 and SUMO-3 but less so for SUMO-1(Gong *et al* 2006; Di Bacco *et al* 2006). The structures of these two proteins are still unknown.

The nucleolar localization of SENP3 suggests that it may regulate certain aspects of nucleolar function (Gong *et al* 2006; Di Bacco *et al* 2006; Nishida *et al* 2000). In addition, SENP3 prefers SUMO2/3 as substrates. These two observations have

narrowed down the number of SENP3 substrates. Indeed, SENP3 has been shown to associate with nucleophosmin NPM1, a crucial factor in ribosome biogenesis (Haindl *et al* 2008). SENP3 deconjugates NPM1-SUMO2 conjugates *in vitro* and counteracts ADP-ribosylation factor-induced modification of NPM1 by SUMO2 *in vivo*. Depletion of SENP3 by siRNA interferes with nucleolar rRNA processing and inhibits the conversion of the 32 S rRNA species to the 28 S form, a phenotype similar to knockdown of NPM1. These results define SENP3 as an essential factor for ribosome biogenesis.

SENP3 is another SUMO protease that is a target for ubiquitin-mediated proteosomal degradation but cellular exposure to reactive oxygen species (ROS) H_2O_2 inhibits SENP3 degradation and causes relocation of SENP3 to the nucleoplasm where the protease is exposed to a different set of SUMO substrates (Huang *et al* 2009; Kuo *et al* 2008). Increased reactive oxygen species have been detected in human prostate tumors and not surprisingly elevated SENP3 levels have been detected in prostate cancer as well as ovarian, lung, rectum and colon carcinomas (Han *et al* 2010).

Similar to SENP1 and SENP2, SENP3 has been implicated in transcription, though its cellular interactions have been studied with less scrutiny. While SENP1 has been known to interact directly with HIF1α through deSUMOylation resulting in subsequent upregulation of HIF1α-dependent genes, SENP3 indirectly influences HIF1α transcription through interaction with the coregulator p300, whose deSUMOyation by SENP3 results in enhanced binding of p300 to HIF1α and similar

upregulation of HIF1α-dependent genes (Huang *et al* 2009).

Using PML SUMOylation mutants as model substrates, SENP5 can remove poly-SUMO-2 or poly-SUMO-3 from the Lys^{160} or Lys^{490} positions of PML. However, SENP5 could not remove SUMO-1 from the Lys^{160} or Lys^{490} positions of PML. Nonetheless, SENP5 could remove SUMO-1, -2, and -3 from the Lys^{65} position of PML. Thus, SENP5 also possesses limited SUMO-1 isopeptidase activity (Gong *et al* 2006).

Knockdown of SENP5 by RNA interference resulted in increased levels of SUMO-1 and SUMO-2/3 conjugates, inhibition of cell proliferation, defects in nuclear morphology, and appearance of binucleate cells, revealing an essential role for SENP5 in mitosis and/or cytokinesis (Di Bacco *et al* 2006).

4. The crosstalk between Ub and SUMO

First, ubiquitination was thought to be merely relegated to the control of protein turnover and degradation, whereas the attachment of SUMO was involved in the regulation of protein activity and function. However, the boundaries between the protein fates related to these tag molecules are becoming more and more fuzzy, as either the differences between mono-, multi- and poly-modifications or the lysine residue used for growth of the poly-chains is being dissected. The Ub and SUMO pathways are no longer separated, and many examples of this cross-talk are found in the literature, involving different cellular processes ranging from DNA repair and genome stability, to the regulation of protein subcellular localization or enzyme activity. The same protein can be conjugated to SUMO and Ub for antagonistic, synergistic or multiple outcomes, illustrating the intricacy of the cellular signaling networks. Ub and SUMO have met and are now applying for new regulatory roles in the cell (Denuc *et al*, 2010).

Signaling through covalent protein modifications requires the recognition by specific effectors of each type of modification. Usually, these effectors are modular domains embedded in larger host proteins, which, on binding to the tagged protein, shift their molecular conformation and trigger a molecular response. These motifs can be present as single or multiple arrayed domains. More than 15 UBDs (Ub-binding domains) (Hurley *et al*, 2006) have been described to date, which are able to discriminate between the different Ub modification states of a particular tagged protein. In contrast, only one SIM (SUMO interacting motif) has been described to

date (Minty *et al*, 2000; Song *et al*, 2004).

To add further complexity, many proteins are multiply modified and this multiplicity of modifications may act in a combinatorial manner (Yang *et al*, 2005). Therefore it is not so surprising to find situations where Ub and SUMO modifications communicate and even sometimes involve the same lysine residue with opposite functions, in what is classically defined as antagonistic modulation. However, the interplay may not be restricted to simple competition, as multisite modification is a means of co-ordinating dynamic regulatory processes. (Figure 10) (Denuc *et al*, 2010).

Protein	Function	SUMOylation	Lysine residue	Ubiquitination	Lysine residue
Cross-talk between SUMO and Ub pathways					
USP25	DUB	Inactivation	Lys^{99}, Lys^{141}	Activation	Lys^{99}
E2-25K	SUMO E2	Inactivation	Lys^{45}		
Mdm2	Ub E3	Stabilization	Lys^{446}	Degradation	Lys^{446}
SUMO and Ub interplay					
PML	Oncoprotein	Favours ubiquitylation	Lys^{160}	Degradation	Lys^{337}, Lys^{380}, Lys^{394}, Lys^{400}, Lys^{401}, Lys^{476}, Lys^{515}
IκBα	Inhibits NF-κB	Stabilization	Lys^{21}	Degradation	Lys^{21}, Lys^{22}
NEMO	IKK regulator	Accumulation in the nucleus	Lys^{277}, Lys^{309}	Translocation to the cytoplasm	Lys^{277}, Lys^{309}
BMAL1	Transcription factor	Favours ubiquitylation	Lys^{259}	Degradation	Lys^{259}
PCNA	Translational factor	Prevention of homologous recombination	Lys^{164}	Mono-Ub: error-prone DNA repair. Poly-Ub: error-free DNA repair	Lys^{164}

Figure 10 - Examples of proteins post-translationally modified by SUMO and Ub (Denuc et al, 2010)

Among the RING-type E3 ligases, RING finger protein 4 (RNF4 or SNURF) is proposed to ubiquitylated proteins that have been modified by SUMO and therefore is classified into a special category of E3s, termed as the SUMO-targeted ubiquitin ligases (STUbLs) (Perry *et al*, 2008; Sriramachandran *et al*, 2014).

The discovery of STUbLs (SUMO-targeted ubiquitin ligases) directly links the SUMOylation process to ubiquitination pathways. By means of tandem SIMs, STUbLs recognize poly-SUMOylated proteins and target them for Lys48-linked polyubiquitylation and degradation through their E3 ubiquitin ligase activities. Only a few STUbLs have been identified thus far, including Slx5–Slx8 (where Slx is synthetic lethal of unknown function) in *Saccharomyces cerevisiae*, Rfp1 (RING finger protein 1)/Rfp2–Slx8 in *Schizosaccharomyces pombe*, human RNF111/Arkadia (Poulsen *et al*, 2013) and RING finger 4 (RNF4) in mammalian cells (Tatham *et al*, 2008). RNF4 is a dimeric STUbL. Its N-terminal half contains four tandem SIM repeats (SIM1–SIM4), referred to collectively in the present study as the SIMs domain, that recognize poly- SUMOylated substrates. The RING domain at the C-terminal half acts together with the SIMs domain to facilitate ubiquitination of proteins modified with poly-SUMO chains (Tatham *et al*, 2008; Plechanovova *et al*, 2012). Bruderer *et al.* (Bruderer *et al*, 2011) identified more than 900 putative endogenous poly-SUMOylated proteins by using the RNF4-SIMs domain as bait. A computational string search identified additional novel SIM clusters in many other proteins (Sun *et al*, 2012). More recently, the human RNF111/Arkadia was identified as a new STUbL, using three adjacent SIMs for specific recognition of poly-SUMO2/3 chains and using Ubc13–Mms2 as a cognate E2 enzyme to promote the Lys63-linked ubiquitination of SUMOylated target proteins (Poulsen *et al*, 2013). RNF4 relies on the SIMs domain for selective binding of poly-SUMO chains over monomer SUMO. For instance, only poly-SUMOylated PML (promyelocytic

leukaemia) proteins can be recognized by RNF4 (Lallemand-Breitenbach *et al*, 2008); however, the structural basis of this recognition is unclear. Recently, Keusekotten *et al.* (Keusekotten *et al*, 2013) showed that the sequence and spacing of the RNF4-SIMs domain regulate the avidity-driven recognition of poly-SUMO chains. They also concluded that the SIM2 and SIM3 regions are necessary and sufficient for binding to a SUMO chain, whereas the SIM4 region is only needed for recognition of longer chains.

Objectives

1. Objectives of USP28 research

1.1 Purification of the N-terminal region of USP28 and USP25 de-ubiquitinases.

1.2 Characterization of the covalent and non-covalent interaction between SUMO isoforms and the N-terminal region of USP28 and USP25.

1.3 USP28 SUMOylation primary site identification of the N-terminal region of USP28 by Mass Spectrometry.

1.4 Crystallization of the N-terminal USP25 and USP28 constructs.

1.5 Purification and proteolytic activity analysis of the Catalytic Domain (CD) constructs of USP28 and USP25.

1.6 Identification of the di-Ubiquitin chain specificity of USP28.

1.7 SUMOylation of the Catalytic Domain (CD) of USP28 and SUMOylation primary site identification by Mass Spectrometry.

1.8 Characterization of the proteolytic activity of the Catalytic Domain constructs of USP28.

1.9 Characterization of the proteolytic activity on the SUMO modification of USP28.

1.10 Crystallization of the USP25 or USP28 de-ubiquitinases.

2. Objectives of SENP5 research

2.1 Structural characterization and purification of the catalytic domains of SENP5 and SENP3.

2.2 Functional characterization of the SENP3 and SENP5 catalytic domains.

2.3 Characterization of the inactive mutant of SENP5 (SENP5 C712S). Complex formation between SENP5 C712S with SUMO2 precursor and with the substrate RanGAP1-SUMO2/3.

2.4 Crystallization of the SENP5 in complex with the SUMO2 precursor and with the RanGAP1-SUMO2 substrate.

Materials and Methods

1. Materials and Methods of USP28 research

1.1 General protein purification method

Proteins were amplified by general PCR program using Phusion polymerase (Table 1) and cloned into the vector pET28b-Smt3 or pET28b to encode a polypeptide fused to a thrombin-cleavable N-terminal hexa-histidine tag and followed by another SENP-cleavable Smt3 tag. Expression constructs were used to transform *E. coli* BL21 (DE3) codon plus cells (Novagen). Bacterial cultures were grown by fermentation at 37 °C to A600=0.6, and isopropyl-β-D-thioga-lactopyranoside (IPTG) was added to a final concentration of 0.5 mM. Cultures were incubated for 4-5 h at 30 °C and harvested by centrifugation (6000g, 20min), and the supernatant was discarded. Cell suspensions were equilibrated in 20% sucrose, 20 mM Tris-HCl (pH 8.0), 1 mM β-mercaptoethanol, 350 mM NaCl, 20 mM imidazole, 0.1% IGEPAL CA-630 and 10mM MgCl2, and cells were disrupted by sonication. Cell debris was removed by centrifugation (40,000g). Protein was separated from lysate by metal affinity chromatography using nickel-nitrilotriacetic acid-agarose resin (Qiagen) and eluted with 20 mM Tris-HCl (pH 8.0), 350 mM NaCl, 300 mM imidazole, and 1 mM β-mercaptoethanol and dialyzed against buffer containing 20 mM Tris-HCl (pH 8.0), 250 mM NaCl, and 1 mM β-mercaptoethanol with SENP2 or Thrombin at a 1:1000 ratio. After SENP or Thrombin cleavage, proteins were separated by gel filtration (Superdex 200 or 75; GE Healthcare). Fractions containing the protein of interest were pooled, diluted to 50 mM NaCl, applied to anion exchange resin (Mono Q; GE Healthcare), and eluted with a NaCl gradient from 0 to 50% of a buffer containing 20 mM Tris-HCl (pH 8.0), 1 M NaCl, and 1 mM β-mercaptoethanol in 15 column

volumes (Optional Step if not necessary). Fractions containing the protein of interest were pooled, concentrated to around 10 mg/ml, and snap-frozen in liquid nitrogen prior to storage at -80 °C.

Program	**Temperature ©**	**Time**
1	95	3 min
2	95	1 min
3	58	0.5 min
4	72	1-2 minutes
5	goto 2, 30 cycles	
6	72	8 min
7	10	hold

Table 1 - General PCR program.

1.2 USP28 and USP25 truncation constructs

The original plasmids pENTR-USP25-FL and pENTRC-USP28-FL are supplied by Professor Travis Stracker from Institute for Research in Biomedicine (IRB Barcelona). The wild type of the USP28 and USP25 constructs: USP28- NT(1-159), USP28-CDΔNT-OLD （160-673）, USP28-CDΔNT-NEW （160-757）,

USP28-CD+NT-OLD(1-673) , USP28-CD+NT-NEW(1-757), USP28-FL （Full Length）, USP25-NT(1-165), USP25-CDΔNT-OLD（165-681）, USP25-CDΔNT-NEW （165-739）, USP25-CD+NT-OLD(1-681) , CD+NT-NEW(1-739) and FL（Full Length）were amplified by PCR and then purified by the procedure mentioned above (Primers are in Table 2).

Primers' Name	Sequence (5'-3')
USP25-BamHI-F	GGATCCATGACCGTGGAGCAGAACGT
USP28-BamHI-F	GGATCCATGACTGCGGAGCTGCAGCA
USP28-NT-Stop-NotI-R	GCGGCCGCTAAACCATCAACTCTCCTCCAGTCA
USP25-Stop-NotI-R	GCGGCCGCTTATCTTCCATCAGCAGGAG
USP28-Stop-NotI-R	GCGGCCGCTTATTTCACTGTCACAGTTG
U28-K64R-forw	CTCACTGATGAGAGAGTTAGGGAGCCCAGTCAAGACACT
U28-K64R-rev	AGTGTCTTGACTGGGCTCCCTAACTCTCTCATCAGTGAG
U28-K115R-forw	AGTCTACTGGAGTCTCCCAGAATTCAAGCTGATGGAAGA
U28-K115R-rev	TCTTCCATCAGCTTGAATTCTGGGAGACTCCAGTAGACT
MUT-U28NT-K99R-F	CCTTACTCATGATAACAGAGAGATCTTCAGGCTG
MUT-U28NT-K99R-R	CAGCCTGAAGATCATCTCT GTT ATC ATG AGT AAG G
U28CD757 ORI-F	CAAACACAGCCCGTGCCTATGAGAAGAGCGGTGTAG
U28CDΔNT-F	ATA GGA TCC GAT GGT TGG CCA GTT GGG CTG
U25FL-R	ATA GCGG CCG TTA TCT TCC ATC AGC AGG AGT TC
U28FL-R	ATA GCGG CCGC TTA TTT CAC TGT CAC AGT TG

U25CD739E-F	GCAGCAGGAGACCCATAATATCTAGAGCAGCCATC
U25CD739E-R	GATGGCTGCTCTAGATATTATGGGTCTCCTGCTGC
U25CD757Y-F	CAAACACAGCCCGTGCCTAAGAGAAGAGCGGTGTAG
U25CD757Y-R	CTACACCGCTCTTCTCTTAGGCACGGGCTGTGTT TG

Table 2 - Primers used in USP28 and USP25 studies

1.3 Mutants of USP28-NT and USP28CD+NT constructs

The mutations of USP28-NT-K64R, USP28-NT-K115R, USP28-NT-K99R, USP28-NT-K64RK115R (DM: Double Mutant) and USP28-NT-K64RK99RK115R(TM: Triple Mutant); USP28-CD+NT-K115R, USP28-CD+NT-K99R, USP28-CD+NT-K99RK115R (DM: Double Mutant) were introduced into the USP28 NT and USP28 CD+NT using QuikChange mutagenesis kit (Stratagene). They were constructed by mutagenic PCR program using Turbo Pfu polymerase (Table 3), treated with DPN1 for 2h, purified by agarose gel and subsequently left with PNK and T4 DNA ligase to produce the circular plasmid. (Primers are in Table 2)

Program	Temperature ©	Time
1	95	3 min
2	95	1 min
3	58	1 min
4	72	7-8 minutes
5	goto 2, 12-18 cycles	
6	10	hold

Table 3 - Mutagenic PCR program.

1.4 SUMO Constructs

Plasmids containing Δ14-SUMO2-Precursor (First 14AA Deletion of SUMO2 precursor) and Δ14-SUMO2 (First 14AA Deletion of the mature SUMO2) were constructed at the Sloan-Kettering Institute in New York by David Reverter. They were purified as mentioned in section 2.1.1 of Methods.

1.5 SUMOylation of USPs (USP28-NT and USP28-CD+NT constructs and mutants)

The small scale SUMOylation reaction of USPs were formed in a reaction mixture containing 20mM Hepes pH7.5, 5mM MgCl2, 0.1% Tween20, 50mM NaCl,

1mM dithiothreitol, 1mM ATP, 150nM SAE1/SAE2 (E1)), 100nM Ubc9 (E2), 10nM IR1 (E3),16mM USPs and 32mM Δ14S2 in MilliQ water. Samples are taken at 0‘, 30‘ and 60‘.

The large scale SUMOylation reaction of U28CD+NT-S2 is as 10 time as the SUMOylation reaction test. Products were verified by SDS-page, followed by an ion exchange chromatography to separate non-modified USP28 from the SUMO conjugated to USP28, then concentrated to 1mg/ml and snap-frozen in liquid nitrogen prior to storage at -80 °C.

1.6 Mass spectrometry

Mass spectrometry experiments were performed in the IRB Barcelona mass spectrometry core facility. Proteins were excised from polyacrylamide gels and subjected to in-gel digests with trypsin, chymotrypsin or both enzymes. Digested peptides were diluted in 1% formic acid (FA). The nano-LC-MS/MS set up was as follows. Samples were loaded on a 180 µm × 2 cm C18 Symmetry trap column (Waters) at a flow rate of 15 µl/min using a nano-Acquity Ultra Performance LCTM chromatographic system (Waters Corp., Milford, MA). Peptides were separated using a C18 analytical column (BEH130™ C18 75 µm × 25 cm, 1.7 µm, Waters Corp.) with a 90 min run, comprising three consecutive steps with linear gradients from 1 to 35% B in 30 min, from 35 to 50% B in 5 min, and from 50 % to 85 % B in 3 min, followed by isocratic elution at 85 % B in 10 min and stabilization to initial conditions (A= 0.1% FA in water, B= 0.1% FA in CH3CN). The column outlet was directly connected to an

Advion TriVersa NanoMate (Advion) fitted on an LTQ-FT Ultra mass spectrometer (Thermo). The mass spectrometer was operated in a data-dependent acquisition (DDA) mode. Survey MS scans were acquired in the FT with the resolution (defined at 400 m/z) set to 100,000. Up to six of the most intense ions per scan were fragmented and detected in the linear ion trap. The ion count target value was 1,000,000 for the survey scan and 50,000 for the MS/MS scan. Target ions already selected for MS/MS were dynamically excluded for 30 s. Spray voltage in the NanoMate source was set to 1.70 kV. Capillary voltage and tube lens on the LTQ-FT were tuned to 40 V and 120 V. Minimal signal required to trigger MS to MS/MS switch was set to 1000 and activation Q was 0.250. The spectrometer was working in positive polarity mode and singly charge state precursors were rejected for fragmentation. At least one blank run before each analysis was performed in order to ensure the absence of cross contamination from previous samples.

A database search was performed with Proteome Discoverer software v1.3 (Thermo) using the Sequest search engine and a custom database, which included N-term sequences of USP28 and USP25. Search parameters included no enzyme restriction, carbamidomethyl in cysteine as static modification and methionine oxidation and +599.266 Da (QQQTGG) in lysine as dynamic modifications. Peptide mass tolerance was 10 ppm and the MS/MS tolerance was 0.8 Da. Peptides with XCorr > 1.1 (z=1), 1.25 (z=2), 1.68 (z=3) were considered as positive identifications.

1.7 Proteolysis activity assay against different poly-Ub chains of USPs

The Human poly-Ub chains (K48, 1-7; K63, 2-7ubs) and the Human poly-Ub chains (K48, K63, 3-7ubs) were purchased from Boston Biochem and the eight types of Di-UBs (Linear, K6, K11, K27, K29, K33, K48, and K63) were purchased from UBPBio company. They were dissolved in the buffer containing 250mM NaCl, 20nM Tris 8.0 and 1mM β-mercaptoethanol and the final concentration is 1mg/ml.

The poly-Ub chains were diluted 10 times to 0.1 mg/ml and mixed with different concentrations (0.5, 5, 50 and 500 nM) of USPs at 37 °C in a buffer containing 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1% Tween 20, and 2 mM dithiothreitol. When the substrate become to Di-Ub, a narrow dilution of the USPs is used: 4, 20, 100 and 500 nM. Reactions were stopped after 30 min with SDS loading buffer and analyzed by gel electrophoresis (PAGE). Proteins were detected by staining with SYPRO (Bio-Rad).

For the time course experiment of USP28CDNT-S2, USP28CDNT and USP28-FL cleavage of Di-UBs, 50nM USP28s was added to the reaction to digest the Di-UBs to check the change of the USP28 proteolysis activity. Samples are taken at 20‘ and 40‘.

1.8 Crystallization

The proteins of different USPs were sprayed into 96-well sitting drop plates, mixed with purchased screen conditions and stored at either 20°C. Optimized plates were made by mixing 0.5-1μL of protein/complex with corresponding buffer and adding to either sitting or hanging drop plates and stored at 20°C. After incubation of several days, bigger crystals were showing up and they could be fished up and then stored in the liquid nitrogen. The X-ray diffraction was carried out at the Alba or ESRF synchrotron.

2. Materials and Methods of SENP5 research

2.1 General protein purification method

Same method as mentioned in section 1.1 of Methods.

2.2 Different Catalytic Domain (CD) constructs of SENP3 AND SENP5

The original plasmids containing SENP3 and SENP5-FL (Full Length) and the plasmids containing SENP3-NΔ92, SENP3-NΔ290, SENP3-NΔ354, SENP5-NΔ475 and SENP5-NΔ535 were constructed at the Sloan-Kettering Institute in New York by David Reverter. The SENP3-NΔ257, SENP3-NΔ377, SENP3-NΔ390, SENP5-NΔ440, SENP5-NΔ559 and SENP5-NΔ568 were generated by PCR (Table 1) amplification of the indicated residues and subcloning into the pET28b vector (Primers are in Table 4).

Primers' Name	Sequence (5'-3')
SENP5 568M pri 1	TAC GCT AGC ATG CTG GAT ATG GAC G
SENP5 568M pri 2	CTC GAA TTC TCA GTC CAT GAG CCG
SENP3 390M pri 1	ATA GCT AGC ATG GAT GAC TTG GGG ACC
SENP3 390M pri 2	CTC GAA TTC TCA CAC AGT GAG TTT GC
SENP3 377 Prime F	ATA GCT AGC GGC TTC CGA GTG GCT TAT AAG C
SENP5 559 Prime F	CGG GCT AGC AAC TTC CGT ATC TTC TAT AAT AAA C
SENP 3 4T1-390 forw	ATA GGA TCC ATG GAT GAC TTG GGG ACC
SENP 5 4T1-440 forw	ATA GGA TCC ATG GAG GAG GAT GGA TCT C
SENP 5 4T1-568 forw	CTC GGA TCC ATG CTG GAT ATG GAC GAC C

SENP 5 28A-440 forw	**ATA GCT AGC ATG GAG GAG GAT GGA TCT C**
S5 FL-4T P1	CGC GGA TCC ATG AAA AAA CAG AGG AAA ATT C
S5 FL-4T P2	ATA CTC GAG TCA GTC CAT GAG CCG GCA CTC
S3 FL-4T P1	CGC GGA TCC ATG AAA GAG ACT ATA CAA GGG
Sp568-712M-P1	GAAAAACGACAGTGACTCTGGAGTCTTTGTGCTCC
Sp568-712M-P2	GGAGCACAAAGACTCCAGAGTCACTGTCGTTTTTC

Table 4 - The primers used in SENP3 AND SENP5 studies

2.3 SUMO Constructs for the proteolysis activity analysis of SENP5-CD

Plasmids containing full-length SUMO1, SUMO2 and SUMO3, SUMO1GG, SUMO2GG and SUMO3GG with an insertion of two additional glycine residues after the Gly-Gly motif were constructed at the Sloan-Kettering Institute in New York by David Reverter. Plasmids were expressed from pET28b in E. coli Bl21 (DE3) codon plus cells (Novagen) and purified by excluding the native stop codon and fusing a C-terminal hexa-hisitidine tag C-terminal to the native polypeptide. SUMOs were purified by Ni-NTA-affinity chromatography (Qiagen) and separated by gel filtration (Superdex 200; GE Healthcare). Fractions containing the protein of interest were pooled, concentrated to 10 mg/ml, and snap-frozen in liquid nitrogen prior to storage at -80 °C.

Plasmids containing Δ18SUMO1 (18-101), SUMO2ΔGG (1-91) and Δ14SUMO2 (15-95) were also constructed at the Sloan-Kettering Institute in New York by David Reverter. Proteins were purified in the same manner as wild type

proteins and concentrated to 10mg/ml and snap-frozen in liquid nitrogen prior to storage at -80 °C.

RanGAP1-SUMO1 and RanGAP1-SUMO2 were formed in a reaction mixture containing 20mM Hepes pH7.5, 5mM MgCl2, 0.1% Tween, 50mM NaCl, 1mM dithiothreitol, 1mM ATP, 150nM SAE1/SAE2 (E1)), 100nM Ubc9 (E2), 16mM NΔ419RanGAP1 and 32mM Δ14SUMO2/Δ18SUMO1 in MilliQ water. DiSUMO2 was formed in a reaction mixture containing 20mM Hepes pH7.5, 5mM MgCl2, 0.1% Tween, 50mM NaCl, 1mM dithiothreitol, 1mM ATP, 150nM SAE1/SAE2 (E1)), 100nM Ubc9 (E2), IR1 (E3), 32mM Δ14SUMO2 and 16mM SUMO2ΔGG in MilliQ water. Products were verified by SDS-page, purified by gel filtration (Superdex 200 or 75), concentrated to 10mg/ml and snap-frozen in liquid nitrogen prior to storage at -80 °C.

2.4 SENP5-CD(568-755) Proteolysis activity assay against different SUMOylated substrates

The full-length pSUMO1, pSUMO2 and pSUMO3, pSUMO1GG, pSUMO2GG and pSUMO3GG with an insertion of two additional glycine residues after the Gly-Gly motif as well as the RanGAP1-SUMO1, RanGAP1-SUMO2 and Di-SUMO2 were purified as mentioned above. They were dissolved in the buffer containing 250mM NaCl, 20nM Tris 8.0 and 1mM β-mercaptoethanol and the final concentration is 1mg/ml.

The pSUMO1, pSUMO2 and pSUMO3, pSUMO1GG, pSUMO2GG and

pSUMO3GG were diluted 10 times to 0.1 mg/ml and mixed with different concentrations (150nM to 0.15nM) of SENP5-CD at 37 °C in a buffer containing 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1% Tween 20, and 1 mM dithiothreitol. When the substrates become to RanGAP1-SUMO1, RanGAP1-SUMO2 and Di-SUMO2, a wide dilution of the SENP5 is used: 150nM to 1.5pM. Reactions were stopped after 25 min with SDS loading buffer and analyzed by gel electrophoresis (PAGE).

For the time course experiment of SENP5-CD cleavage of RanGAP1-SUMO1, RanGAP1-SUMO2, and 0.5nM SENP5-CD was added to the reaction to digest the RanGAP1-SUMO1, RanGAP1-SUMO2 to check the USP28 proteolysis activity. Samples are taken at 5‘, 10‘, 20‘ and 40‘.

2.5 The inactive mutant of SENP5-CD (568-755)-C712S:

The mutant SENP5-CD (568-755)-C712S was introduced into the SENP5-CD (568-755) using QuikChange mutagenesis kit (Stratagene). It was constructed by PCR (Table 3), treated with DPN1, purified by agarose gel and subsequently left with PNK and T4 DNA ligase to produce the circular plasmid. (Primers are in Table 4). The same proteolysis activity assay was made of this SENP5-CD (568-755)-C712S with RanGAP1-SUMO to verify it is inactive.

2.6 The SENP5-CD-C712S and SUMO2-precursor/RanGAP1-S2 complex formation and crystallization:

The inactive mutant of catalytic domain of SENP5 (SENP5C712S) was concentrated to >10mg/ml and added to a reaction mixture containing 5mM NaCl, 20mM Tris, pH 8.0 and 1mM β-mercaptoethanol. Complexes were attempted by adding a 2:1 dilution of the inactive protease to the substrate-human SUMO2-precursor and the RanGAP1-SUMO2 substrate. The mixtures were purified by gel filtration (Superdex 200; GE Healthcare) and fractions containing the complex were verified by SDS-PAGE and pooled to a final protein concentration of >10mg/ml. The complexes were sprayed into 96-well sitting drop plates, mixed with purchased screen conditions and stored at either 20°C. Optimized plates were made by mixing 0.5-1μL of protein/complex with corresponding buffer and adding to either sitting or hanging drop plates and stored at 20°C. After inoculating several days, bigger crystals were showing up and they could be fished up and then stored in the liquid nitrogen. The X-ray diffraction was carried out at the Alba or ESRF synchrotron.

Results

1. The Results of USP part

1.1 Structural and functional characterization of USP28 and USP25

Based on structural alignments with other members of the USP family, and on a previously published report on the homologous USP25, both USP25 and USP28 full-length protein (1055 and 1077 residues) can be divided in three domains: the N-terminal domain (around 160 residues long), the catalytic –conserved" USP domain (around 450 – 500 residues long) and the C-terminal extension domain (around 300 residues) (Figure 1).

In a recent report on USP25, *in silico* comparative searches predicted three different ubiquitin binding motifs in the N-terminal region: one UBiquitin Associated Domain (UBA) and two Ubiquitin Interacting Motifs (UIM) (Denuc *et al*, 2009). The N-terminal regions of USP25 and USP28 are highly homologous, particularly in these ubiquitin binding motifs (Figure 1).

It is also reported that USP25 is a target for SUMO conjugation that is preferentially conjugated with SUMO2/3 as a consequence of its favorable binding to SUMO2/3, compared to SUMO1. Two SUMOylation sites are identified within the first and directly adjacent to the second ubiquitin interaction motif. These UIMs are required for the full catalytic potential of USP25 toward ubiquitin chains. Importantly, SUMOylation of USP25 impairs its activity as a result of reduced affinity to ubiquitin chains (Meulmeester *et al*, 2008).

Figure 1 - Sequence alignment of USP25 and USP28. The N terminal (NT) is about 160 AA and the Catalytic Domain (CD) is about 450-500 AA. The red mark of CD-End shows different ends we used in our study. The K99 is the primary SUMOylation site we have identified. The R230-N649 of USP28 and the R237-N656 of USP25 belonging to the catalytic domain are not shown. The V769-I1068 of USP28 and the H752-L1047 of USP25 belonging to the C-terminal extension domain are not shown.

Recently, the NMR structure of the USP28 N-terminal region (PDB code 2LVA) was deposited by the Northeast Structural Genomics Consortium (Lemak *et al*, 2012) (Figure 2). The NMR structure confirms the presence of the predicted UBA domain in USP28, from Gln22 to Lys64, forming a characteristic 3-helix bundle domain, and constituting the only globular domain in the N-terminal region of USP28 (Figure 1 and 2). The other regions of the N-terminal domain are disordered, with the exception of the formation of an α-helix, from Asp100 to Ser113, which would correspond to one of the predicted UIM domains in USP25 (Figure 1 and 2) (Denuc *et al*, 2009). The second predicted UIM domain, displaying a lower level of homology in USP28, only forms a short 3_{10}-helix in the NMR structure of USP28, from Arg121 to Leu123 (Figure 1 and 2). Thus the NMR structure of the N-terminal region USP28 suggests that it is mainly disordered, with only a few secondary structure elements forming the 3-helix bundle UBA domain and an isolated α-helix corresponding to the first predicted UIM domain. In addition to the presence of these ubiquitin-binding motifs elements, in USP25 it was described the presence of a SUMO-Interacting Motif (SIM) (Meulmeester *et al*, 2008), which is conserved in USP28 and corresponds to the region from Val91 to Leu94 (Figure 1). In USP25, this SIM was elegantly described to participate in the SUMO conjugation reaction by using a novel conjugation mechanism (Meulmeester *et al*, 2008).

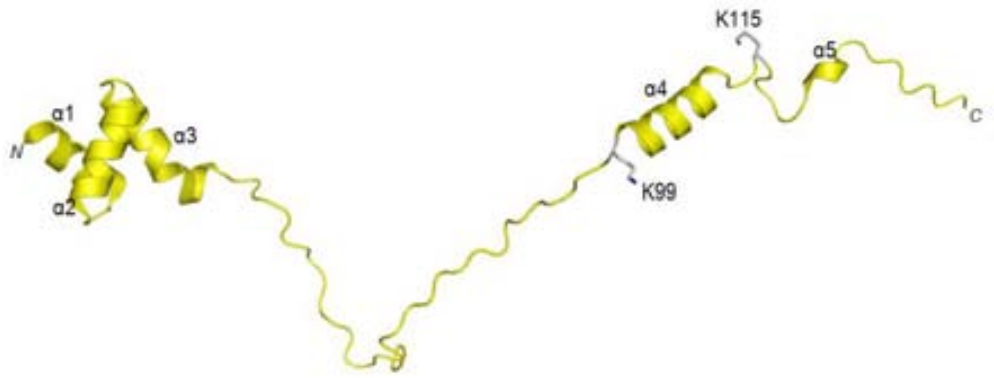
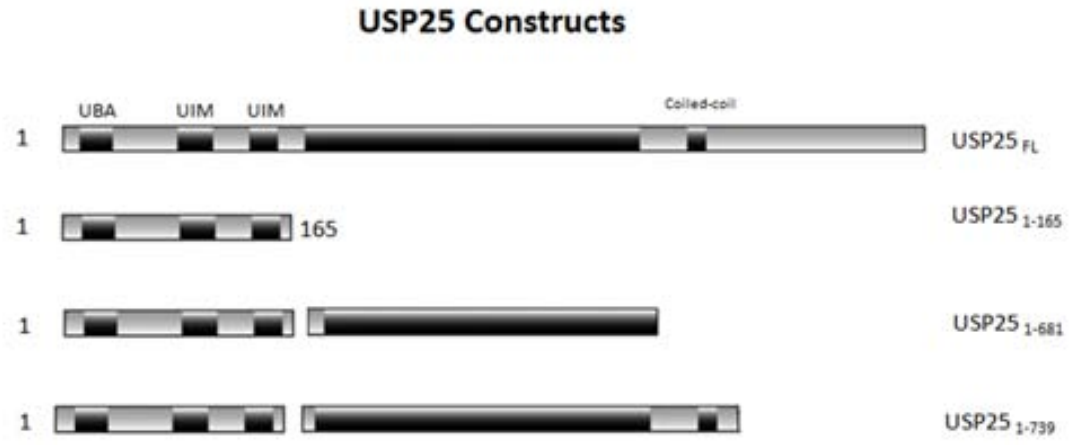


Figure 2 - NMR structure of the N-terminal domain of USP28. Ribbon representation of the deposited NMR structure (PDB code 2LVA) of the N-terminal region of USP28 (residues 22-132). Lys99 and Lys115 side-chains are labeled and shown in stick representation. Secondary structure elements are labeled.

To examine the relative importance of USP25 and USP28's domains to enzymatic activity, we produced different truncation variants of USP25 and USP28 in *E.coli*: USP25 $_{FL}$, USP25 $_{1-165}$, USP25 $_{1-681}$, and USP25 $_{1-739}$; USP28 $_{FL}$, USP28 $_{1-159}$, USP28 $_{1-673}$, USP28 $_{1-757}$, USP28 $_{160-673}$, and USP28 $_{160-757}$ (the sub-index indicates the first and last residue of the construct) (Figure 3).



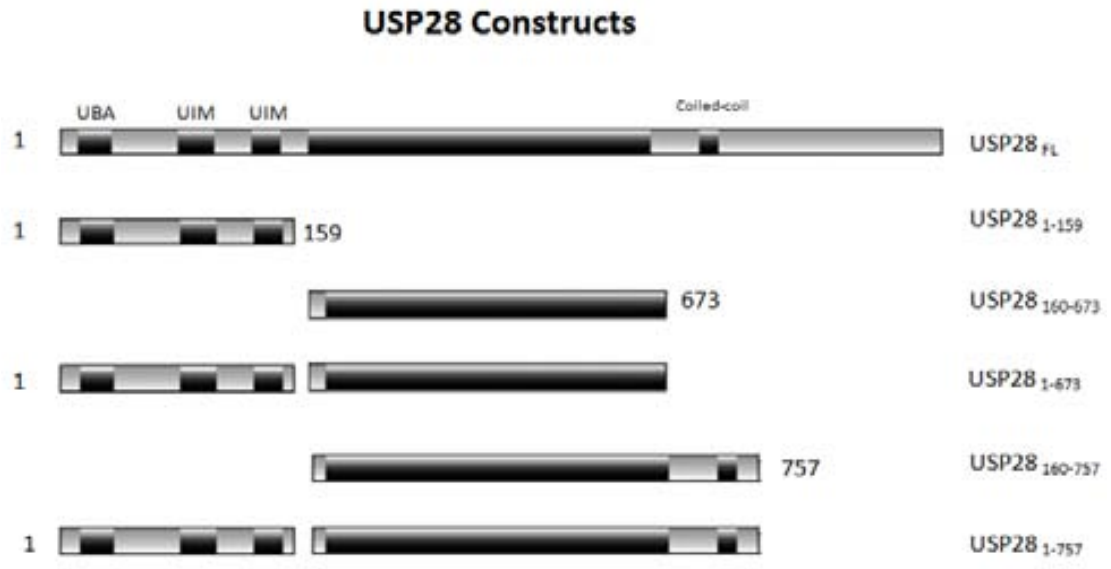


Figure 3 – Constructs of USP25 and USP28 in our study. The N terminal contains the predicted motifs of the UBA and UIMs. The Coiled-coil region is supposed to be a crucial motif to maintain the proteolytic activity.

1.2 Purification of the N-terminal region of USP28 and USP25

Based on the general purification method described in the Methods section, the protein of the N-terminal region of USP25 $_{1\text{-}165}$ and USP28 $_{1\text{-}159}$ can be produced at a high level in *E. coli*. After the SENP2 cleavage of the Smt3 fusion tag, USP28 $_{1\text{-}159}$ can be directly purified by Gel filtration (Superdex 75; GE Healthcare). (Figure 4)

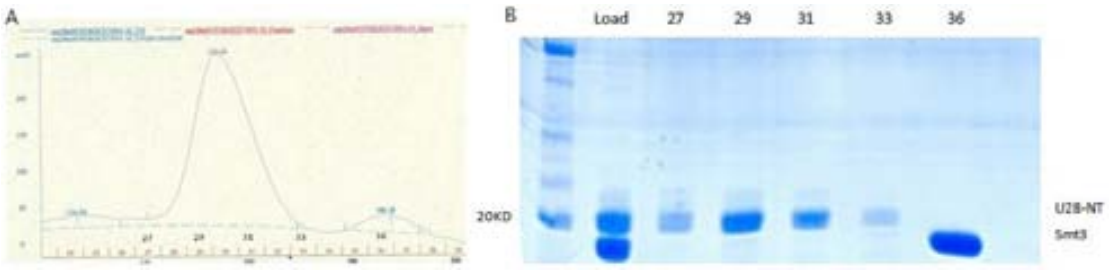


Figure 4 – The purification of USP28 $_{1\text{-}159}$. A. The gel filtration chromatography of USP28 $_{1\text{-}159}$: the peak containing fragments from 27 to 33 is the USP28 $_{1\text{-}159}$. B. The SDS-PAGE of the Gel filtration.

However, when we try to purify the USP25 $_{1\text{-}165}$ by Gel filtration (Superdex 75; GE Healthcare), after the SENP2 cleavage of the Smt3 fusion tag, the Smt3 tag is still mixed with USP25 $_{1\text{-}165}$ (Figure 5A). So, we next made a second purification step through an anion exchange resin (Mono Q; GE Healthcare) and the N-terminal USP25 domain could be finally separated from the Smt3 tag. (Figure 5B and C).

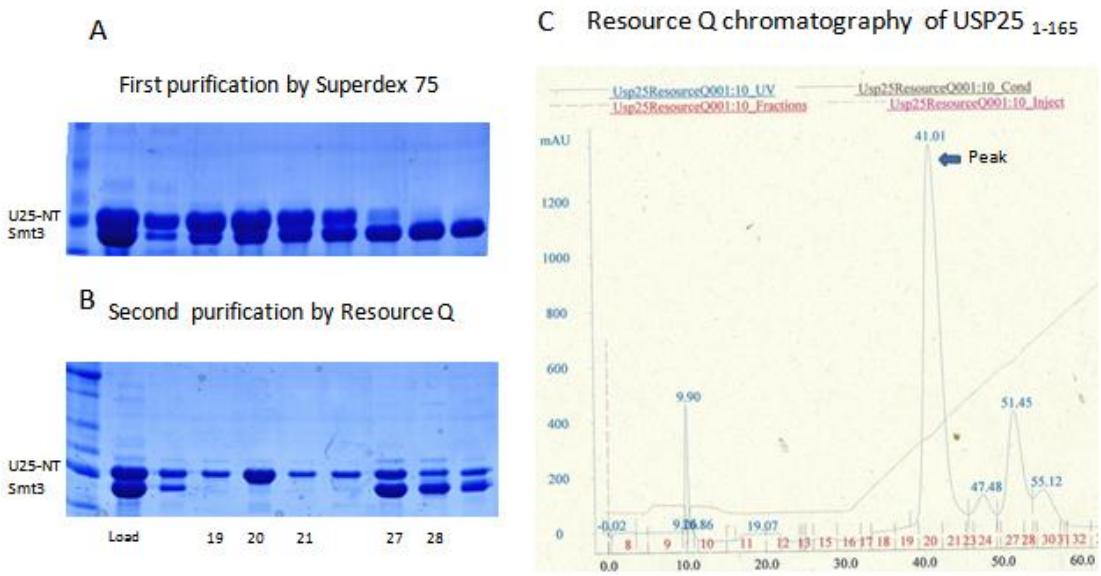


Figure 5 – The purification of USP25$_{1\text{-}165}$. A. After the first purification by Superdex 75, the USP25$_{1\text{-}165}$ is still mixed with the Smt3 tag. B. After the second purification of the anion exchange resin (Mono Q; GE Healthcare), the USP25$_{1\text{-}165}$ elutes separately in a single peak of the fragments 20-21. C. The SDS-PAGE gel of the second purification of the anion exchange resin (Mono Q; GE Healthcare).

Based on these different gel filtration purification processes of the N-terminal regions of USP25 and USP28, they could indicate that probably the USP28-NT is forming a dimer, while the USP25-NT is forming a monomer.

1.3 Covalent and non-covalent interaction test between SUMO and N-terminal region of USP28 and USP25

As mentioned in the Methods section, *in vitro* SUMO covalent conjugation reaction using purified E1, E2 and IR1 (E3 SUMO ligase) resulted in efficient attachment of SUMO1 and SUMO2 to $USP28_{1-159}$ and $USP25_{1-165}$ (Figure 6 and 7). We found that in all cases, at least *in vitro* reaction, the E3 (IR1) is always necessary.

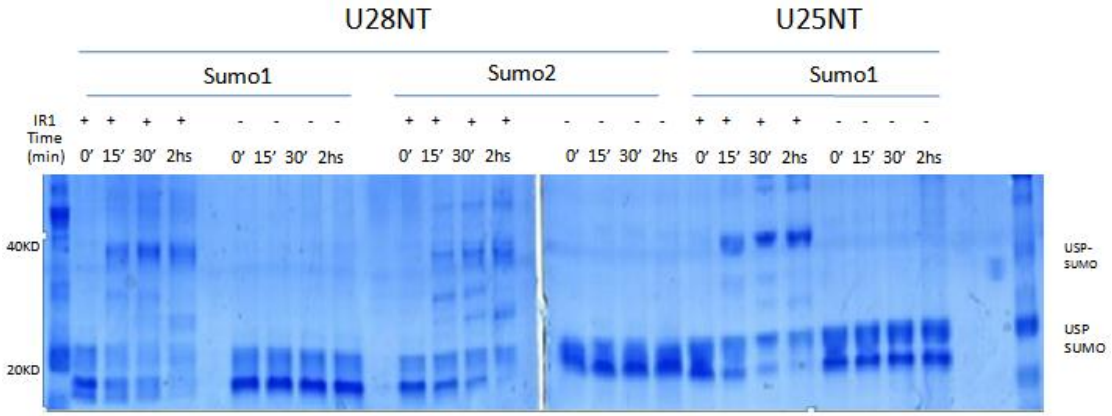


Figure 6 – The SUMOylation of N-terminal region of USP28 and part of USP25. In the case of USP28NT, both SUMOylations with SUMO1 and SUMO2 require IR1 as well as in the case of USP25NT sumoylation with SUMO1.

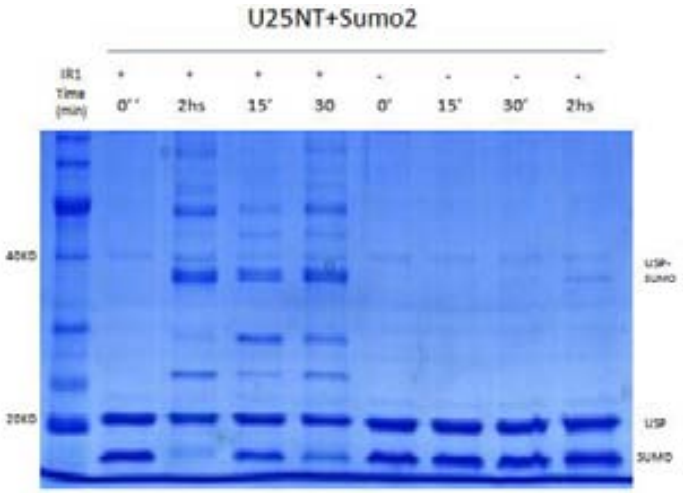


Figure 7 – The SUMOylation of N-terminal region of USP25 with SUMO2. In the case of U25NT, the SUMOylations with SUMO2 also required IR1.

For USP25, it was described that a covalent and a non-covalent interaction of SUMO with the N-terminal region of the protease resulted in an inhibition of de-ubiquitinating activity of USP25 (Meulmeester *et al*, 2008). Here, we also want to check this result.

Therefore, the complex between $USP25_{1-165}$ with SUMO1, $USP28_{1-159}$ with SUMO1 and $USP28_{1-159}$ with SUMO2 have been added to the buffer containing 100mM NaCl, 20mM Tris, pH 8.0 and 1mM β-mercaptoethanol. Complexes were attempted by adding a 2:1 dilution of USP-NTs to the SUMO isoforms. After incubating the complex at room temperature for half an hour, we passed the complexes through a Gel Filtration chromatography (Superdex 200; GE Healthcare). However, we were not able to observe the non-covalent interaction between USP-NTs domains and SUMOs in the elution profiles. In all cases, USP-NTs domains and SUMOs elute separately through the Gel filtration chromatography and they did not seem to interact with each other (Superdex 200; GE Healthcare) (Figure 8).

This result may indicate that there is not non-covalent interaction between USP-NTs domains and SUMOs or that this non-covalent interaction is so weak that cannot be observed by Gel Filtration. We will also try to check this interaction in the section of the USP catalytic domain characterization.

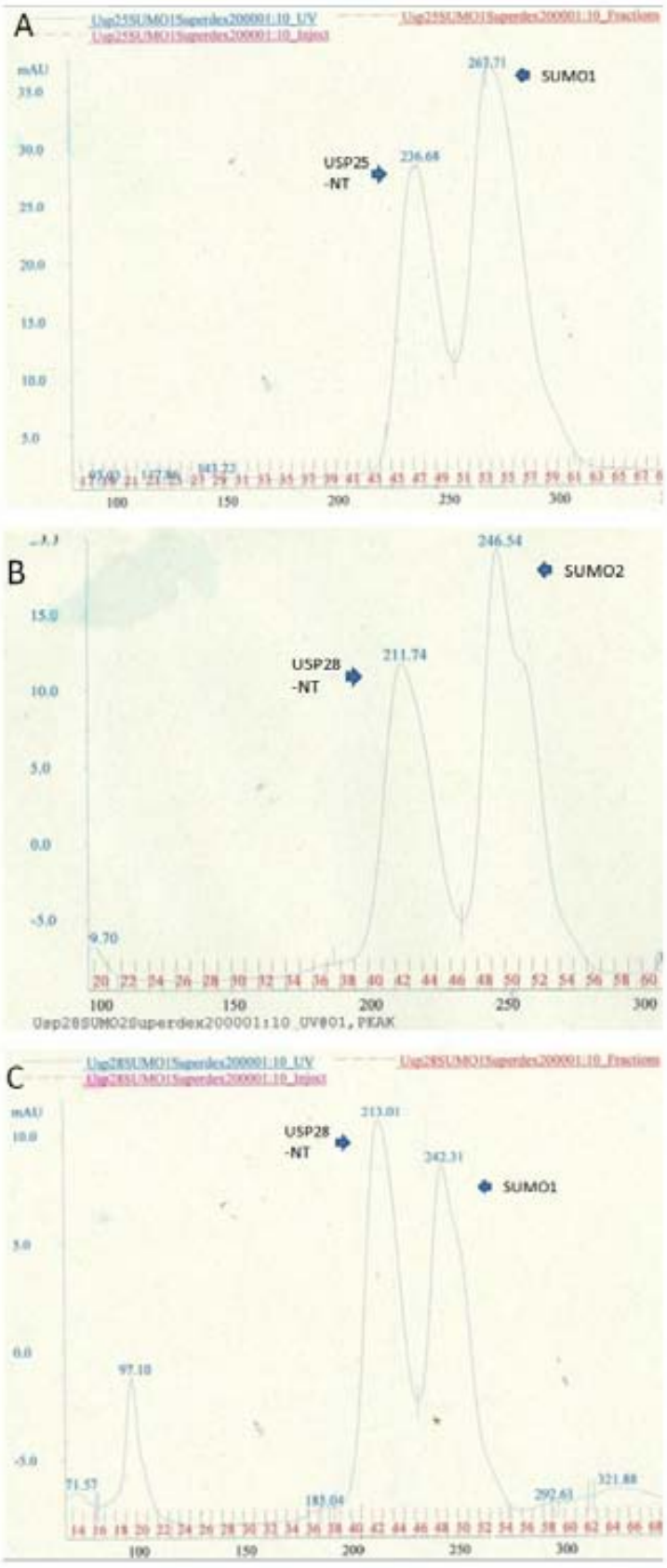


Figure 8 – The non-covalent interaction test of N-terminal region of USP28 and USP25 with SUMO.

A, U25-NT with SUMO1; B, U28-NT with SUMO1; C, U28-NT with SUMO2.

1.4 USP28 SUMOylation primary site identification.

There are already some studies on the USP25 SUMOylation while for the USP28 SUMOylation has not been studied, so we focus on the SUMOylation of the USP28-NT domain. As described in the Methods section, we could successfully set up a SUMOylation reaction of $USP28_{1-159}$ with SUMO2, next we cut the SUMOylated $USP28_{1-159}$ band from the gel and sent it to the Mass Spectrometry facility of the IRB in Barcelona. Mass Spectrometry (MS) analysis of this SUMO-modified USP28 revealed Lys99 as the primary SUMOylation site in the N-terminal domain, followed to a lesser degree by modification on Lys64, Lys85, Lys115 and Lys135 (see Table 1). Remarkably, Lys99 was also found to correspond to the primary site for SUMOylation and mono-ubiquitination in the previous works on USP25 (Meulmeester *et al*, 2008; Denuc *et al*, 2009).

Residue	Non-modified (PSM#)	SUMOylated (PSM#)	Motif ψKXE
K64	12	1	VKEP
K85	198	4	NKEV
K99	177	257	NKDaD
K115	3	7	PKIQ
K135	67	23	TKRS
K138	0	1	RKRK
K140	0	1	RKRC

Table 1 - The Mass Spectrometry analysis of SUMOylated USP28-NT. Sites identified in SUMOylated fragments of the USP28 $_{1-159}$ protein by mass spectrometry. Number of peptide spectrum matches (PSMs) for the non-SUMOylated (left middle column) or SUMOylated (right middle column). Motif compared to linear consensus motif is shown for each (far right column).

To confirm the MS results, we produced single (K99R), double (K115R and K64R), and triple (K115R, K64R and K99R) point mutant constructs in the USP28 N-terminal domain. The SUMO conjugation reaction was conducted using two different E3 ligases, IR1 and Nse2, and despite variable amounts of SUMO conjugation, both ligases confirmed Lys99 as the primary residue for SUMO conjugation in the N-terminal region of USP28 (Figure 9A and 9B). Whereas the USP28 double mutant (K115R and K64R) was conjugated to a similar degree as the wild-type form, the addition of K99R in the USP28 triple mutant strongly decreases conjugation to levels comparable to the K99R single mutant. Interestingly, a SUMO conjugation reaction with a single (K99R) and triple mutant (K115R, K64R and K99R) yielded a faint band in the gel, in a slightly different position as the Lys99 conjugate, probably indicating residual SUMO conjugation on another lysine residue (see asterisks in Figure 9A & 9B).

Based on the deposited NMR structure of the N-terminal domain of USP28 (PDB code 2LVA), Lys99 is located at the beginning of the α-helix corresponding to the first predicted UIM domain. Although Lys99 is not located in a linear SUMO consensus motif (ΨKxE), two reasons might favor this lysine as the major conjugation site: the spatial conformation of Lys99 (together with other E2 interacting residues) in the IUM α-helix, as observed in examples of SUMO conjugation on

lysines located in non-consensus regions (Pichler *et al*, 2005); and the SUMO conjugation enhancement produced by the interaction with SUMO E3 ligases, in our case with IR1 and Nse2. In the case of the homologous USP25, the nearby SIM region was proposed to enhance SUMO conjugation by favoring interaction with the charged E2-SUMO thioester. In our *in vitro* assays, SUMO conjugation of USP28 in the absence of E3 ligases was difficult to observe, although we cannot discard that a similar conjugation mechanism occurs with USP28 as the one described for USP25 (Meulmeester *et al*, 2008).

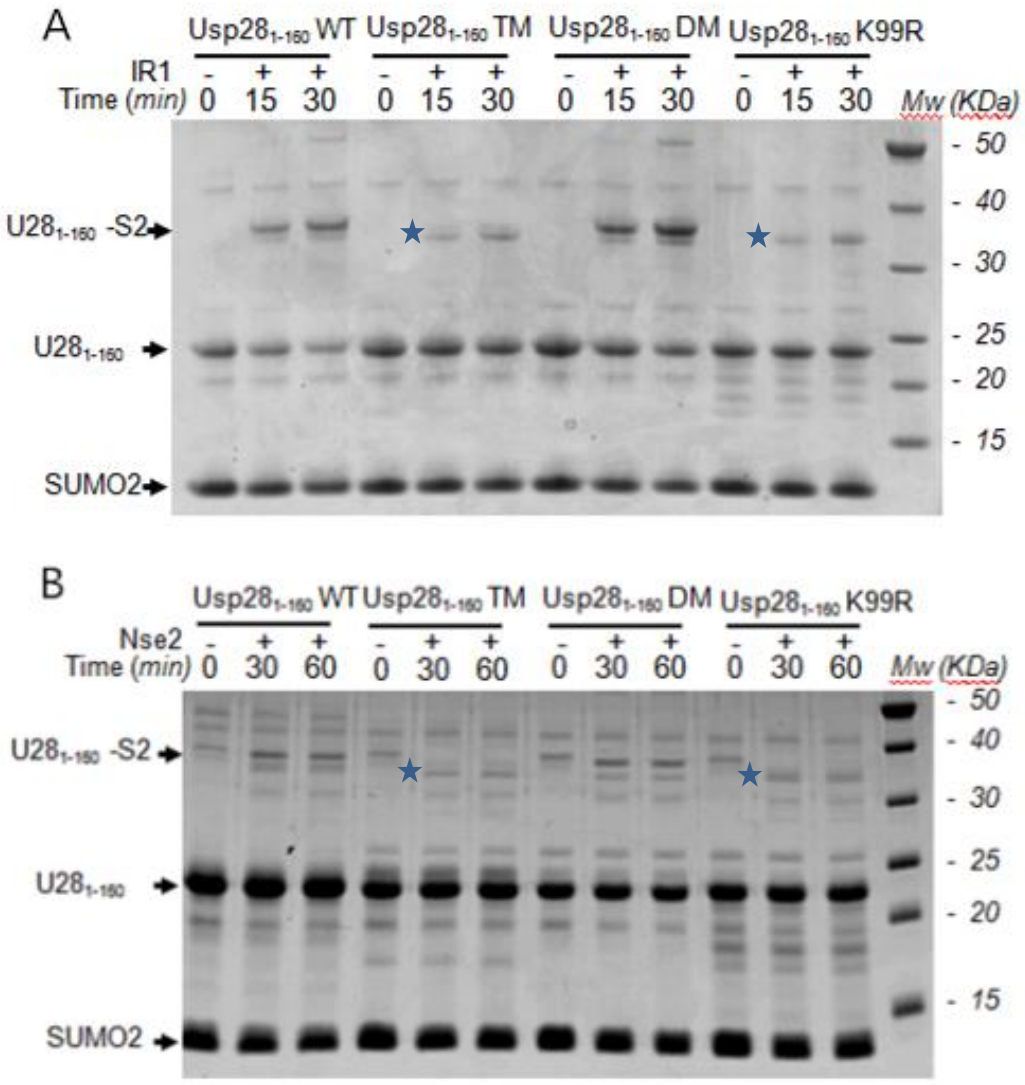


Figure 9 - SUMO conjugation reaction with USP28-NT-Mutants. A. Time course SUMO conjugation reaction with IR1 E3 ligase using different point mutants of the N-terminal region construct of USP28 (USP28 $_{1-159}$). B. Similar reaction as in A but using Nse2 as SUMO E3 ligase.

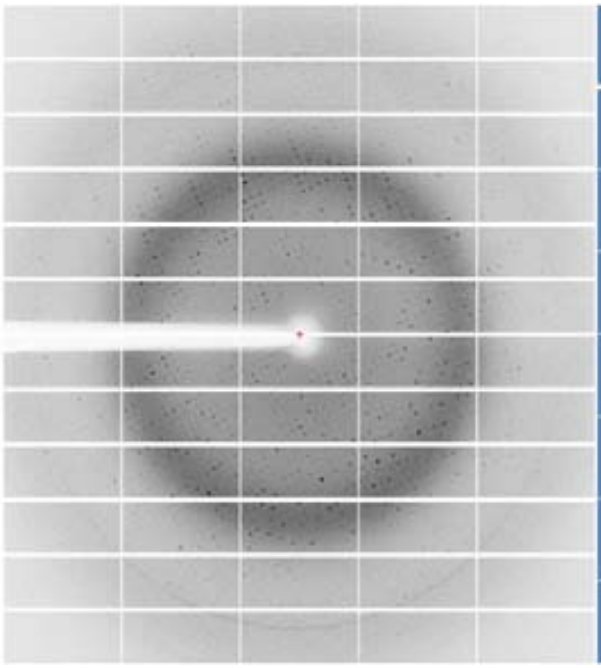
1.5 Crystallization of USP25 and USP28 N-Terminal constructs

Since we have purified the N-terminal region of USP28 and USP25 (USP25 $_{1-165}$ and USP28 $_{1-159}$) in very good yields, we next set up trays with crystallographic robot to try to form protein crystals. Both individual proteases were sprayed into 96-well sitting drop plates, mixed with commercial screen conditions (Molecular Dimension Company) and stored at 18°C to grow crystals. We only found crystals of USP28$_{1-159}$ showing up in a few days in some conditions (Table 2), but we could not find any crystals of USP25 $_{1-165.}$ The next step is the optimization in 24-well-plates, which were made by mixing 0.5-1μL of USP28$_{1-159}$ with corresponding buffer and adding to hanging drop plates and stored at 18°C. After incubation for several days, bigger crystals showed up and were fished up and stored in liquid nitrogen. The X-ray diffraction was carried out at the Alba synchrotron. Unfortunately, the crystals were very anisotropic, displaying only a strong diffracting signal from only one direction (angle), which makes impossible the resolution of the structure. (Diffraction data in Table 3)

Screens	Conditions	Showing up
Proplex	1.5M $AmSO_4$, 0.1M Tris-H 8.0	3rd Day
JCSG Plus	1.6M Tri-Sodium Citrate 6.5	3rd Day

SS1&2	0.2M $MgCl_2$, 30% PEG4000, 0.1M Tris-H 8.5	3rd Day
Morpheus	0.06M Divalents, 0.1M Sys1 PH 6.5, 37.5% MPD_P1K_P3350	3rd Day
Morpheus	0.06M Divalents, 0.1M Sys3 PH 8.5, 30% P550MME_P20K	3rd Day
Morpheus	0.09M NPS, 0.1M Sys3 PH 8.5, 30% P550MME_P20K	3rd Day
Morpheus	0.1M Amino acids, 0.1M Sys1 PH 6.5,30% P550MME_P20K	3rd Day

Table 2 - The screens and conditions of $USP28_{1-159}$ crystals.



Data Collection	U28-NT
Space group	P2
a, b, c (Å)	72.11, 73.52, 88.18
α, β, γ (°)	90, 90.26, 90
Resolution (Å)*	50 – 2.5 (2.18– 2.85)
R_{merge}*	No
I/σI	No
Completeness (%)	No

Table 3 - $USP28_{1-159}$ crystal diffraction pattern. The diffraction of USP28-NT is quite good in one direction, unfortunately due to the anisotropy in the opposite direction we cannot resolve the structure of USP28-NT.

1.6 Purification and proteolytic activity analysis of the Catalytic Domain (CD) constructs of USP28 and USP25

1.6.1 Purification and proteolytic activity analysis of the CD+NT of USP28 (USP28 $_{1-673}$), the CD of USP28 (USP28 $_{160-673}$), the CD+NT of USP25 (USP25 $_{1-681}$ and USP25 $_{1-739}$) and the FL of USP25.

Since we didn‘t observe the non-covalent interaction between the NT of USP28 and USP25 with SUMO, we next try to purify different parts of the Catalytic Domain (CD) of USP28 and USP25 and check the interaction with SUMO as well as the proteolytic activity of these constructs. At the very beginning, the Catalytic Domain we decide to use ranged from the first residue to Leu673 of USP28, and from the first residue to Leu681 of USP25, which we supposed that it corresponded to the functional Catalytic Domain by structural comparisons with other USP family members.

Based on the general purification method described in the Methods section, we first purified the CD+NT of USP28 (USP28 $_{1-673}$) by the Gel filtration (Superdex 75; GE Healthcare) and the CD of USP28 (USP28 $_{160-673}$) by the Gel filtration (Superdex 200; GE Healthcare). After the SENP2 cleavage of the Smt3 fusion tag (Figure 10B and 10D), the SDS PAGE gel shows a molecular weight of USP28 $_{1-673}$ about 75KDa and of the USP28 $_{160-673}$ about 65KDa in (Figure 10 A and C).

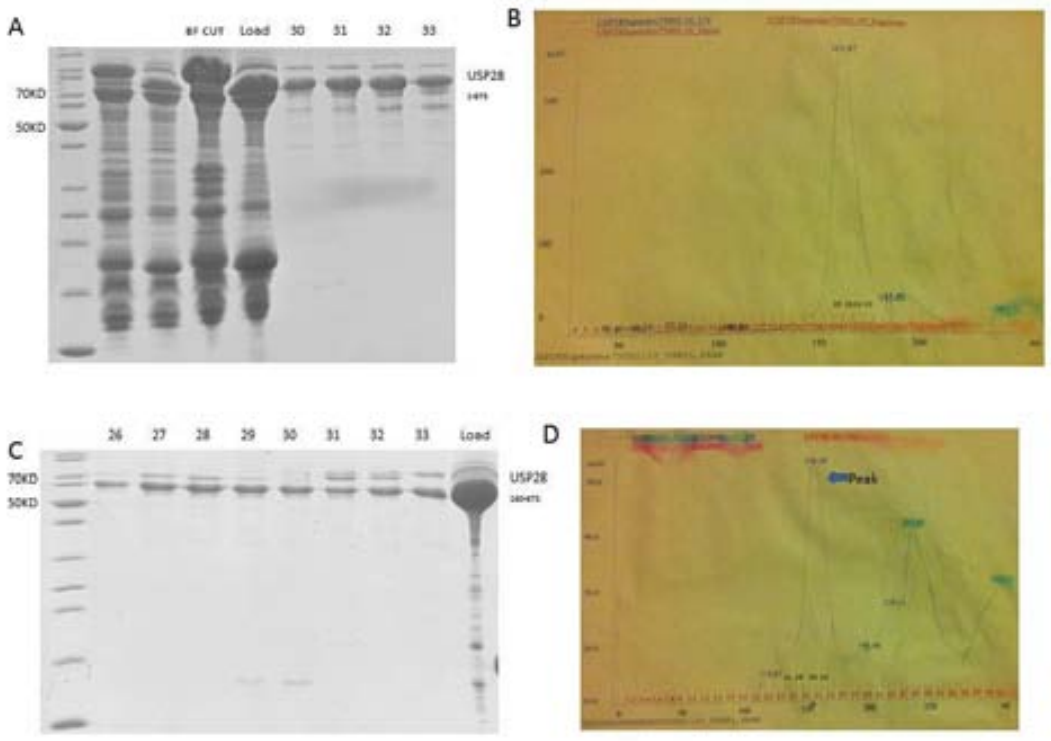


Figure10 – The purification of USP28 $_{1-673}$ and USP28 $_{160-673}$. A. The SDS-PAGE gel of the USP28$_{1-673}$ gel filtration: bf cut: sample before the SENP cleavage of the Smt3 fusion tag, Load: the load sample of gel filtration; B. The gel filtration chromatography of USP28 $_{1-673:}$ the peak containing fragments from 30 to 33 is the USP28 $_{1-159}$; C. The SDS-PAGE gel of the USP28$_{160-673}$ gel filtration: Load: the load sample of gel filtration; D. The gel filtration chromatography of USP28 $_{160-673:}$ the peak containing fragments from 28 to 31 is the USP28 $_{160-673}$.

In order to compare the proteolytic activities between USP28 and USP25 constructs, we also produced the Catalytic Domain (CD) and the Full Length (FL) constructs of USP25. Similarly to the USP28, first we tried to purify the CD+NT of USP25 (USP25 $_{1-681}$) but the protein was not expressed. Then we found in a work from Denuc and coworkers (Denuc *et al*, 2009) explaining ~~"~~USP25m catalytic activity did not strictly depend on the UBDs, but required a coiled-coil stretch between amino acids 679 to 769." This prompted us to design another construct of CD+NT of USP25

(USP25 $_{1\text{-}739}$), which included the coiled-coil stretch between amino acids 679 to 769 (Figure 3). As expected, finally we could purify the USP25 $_{1\text{-}739}$ and the USP25-FL by the Gel filtration (Superdex 200; GE Healthcare) (Figure 11 B and D). The USP25 $_{1\text{-}739}$ is about 85KDa and the USP25-FL is about 120KDa in the SDS PAGE gel (Figure 11 A and C).

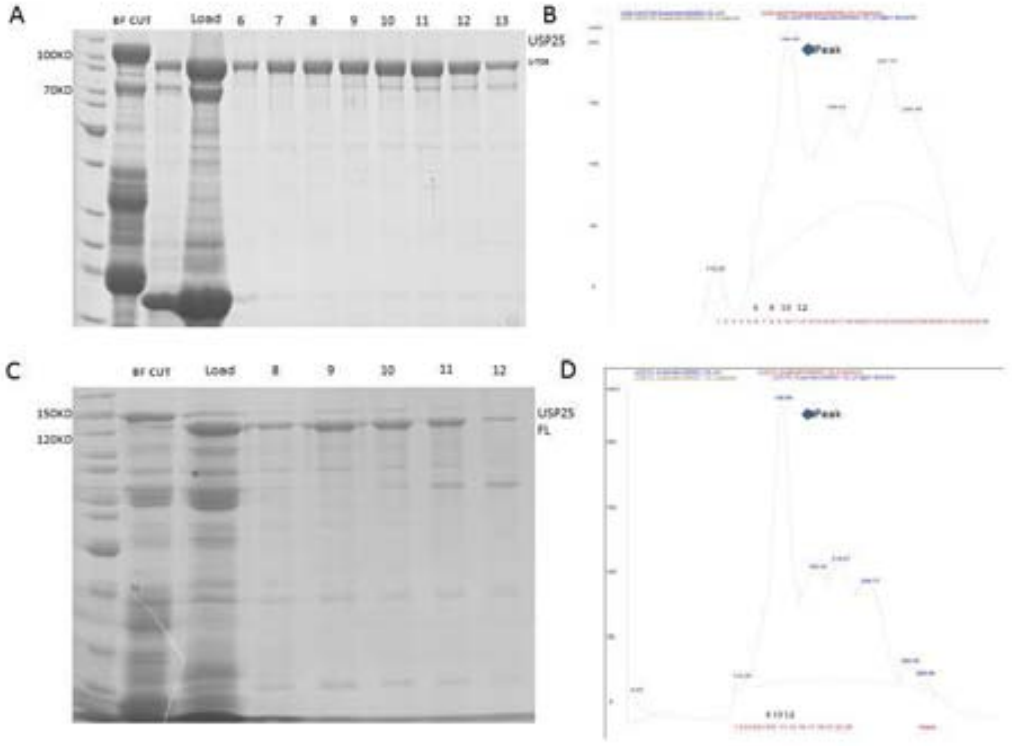


Figure11 – The purification of USP25 $_{1\text{-}739}$ and USP25-FL. A. The SDS-PAGE gel of the USP25$_{1\text{-}739}$ gel filtration: bf cut: sample before the SENP cleavage of the Smt3 fusion tag, Load: the load sample of gel filtration; B. The gel filtration chromatography of USP25 $_{1\text{-}739}$: the peak containing fragments from 8 to 12 is the USP25 $_{1\text{-}739}$; C. The SDS-PAGE gel of the USP25-FL gel filtration: bf cut: sample before the SENP cleavage of the Smt3 fusion tag, Load: the load sample of gel filtration; D. The gel filtration chromatography of USP25-FL: the peak containing fragments from 8 to 11 is the USP25-FL.

With these different constructs of USP25 and USP28, we next carry out the proteolytic activity analysis on two different Ubiquitin substrates: poly UB chain

(K48, 1-7) and poly UB chain (K63, 2-7). We found that despite the CD+NT of USP28 (USP28 $_{1-673}$) construct has the ability to digest the poly UB chain, when compared with the USP25 $_{1-739}$ and USP25-FL, the activity of USP28 $_{1-673}$ is quite low (Figure 12A & B).

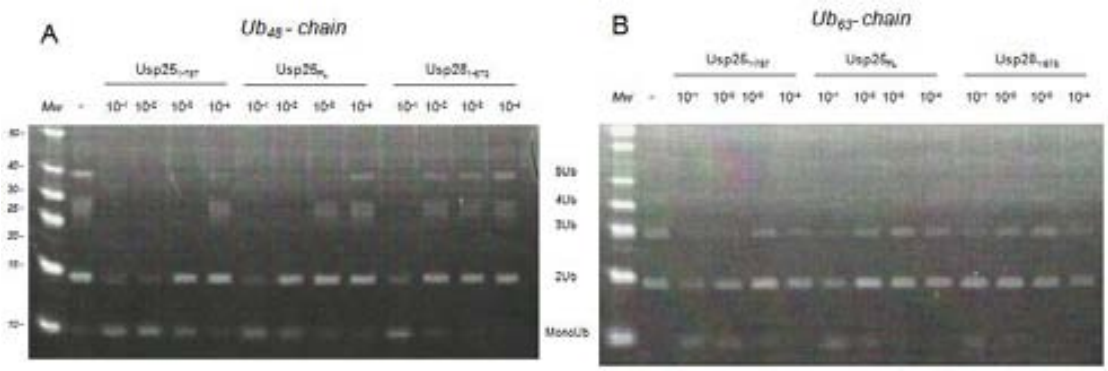


Figure 12 – De-ubiquitinating activity of USP28 $_{1-673}$, USP25 $_{1-739}$ and USP25-FL on K48 and K63 Ubiquitin chain substrates. A. The substrate of Ub chain (K48, 1-7) is the K48 Ub chain mixture containing the mono Ub, 2Ubs, 3Ubs, 4Ubs and 5Ubs. B. The substrate of Ub chain (K63, 2-7) is the K48 Ub chain mixture containing the 2Ubs, 3Ubs, 4Ubs and 5Ubs.

Then we added the CD of USP28 (USP28 $_{160-673}$) to check the proteolytic activity on poly UB chain (K48, 2-7) substrates. We can see that compared with the CD+NT of USP25 (USP25 $_{1-739}$), the CD+NT of USP28 (USP28 $_{1-673}$) and the CD of USP28 (USP28 $_{160-673}$) are mostly inactive (Figure 13). These results may indicate that similarly to USP25, the C-terminal extension of the USP28 catalytic domain is probably required to maintain the proteolytic activity of the catalytic domain.

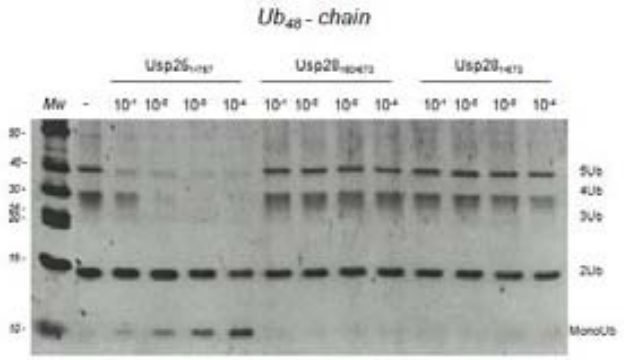


Figure13 –De-ubiquitinating activity of USP25*$_{1-739}$, USP28 $_{1-673}$ *and USP28*$_{160-673}$ *on K48 and K63 Ubiquitin chain substrates. The substrate of Ub chain (K48, 1-7) is the K48 Ub chain mixture containing the mono Ub, 2Ubs, 3Ubs, 4Ubs and 5Ubs.

1.6.2 Purification and proteolytic activity analysis of the new CD+NT of USP28 (USP28 $_{1-757}$), the new CD of USP28 (USP25 $_{160-757}$) and the FL of USP28.

To check these differences in the length of the catalytic domains, we designed new primers to form new constructs of the USP28 containing an extended C-terminal tail. So we purify the new CD+NT of USP28 (USP28 $_{1-757}$), the new CD of USP28 (USP28 $_{160-757}$) and the FL of USP28 constructs (Figure 14) by the Gel filtration (Superdex 200; GE Healthcare) (Figure 14 B, D and F). After SENP2 cleavage, the USP28 $_{1-757}$ is nearly 85KDa, the USP28 $_{160-757}$ is about 75KDa and the USP28-FL is about 120KDa in the SDS-PAGE gel (Figure 14 A, C and E).

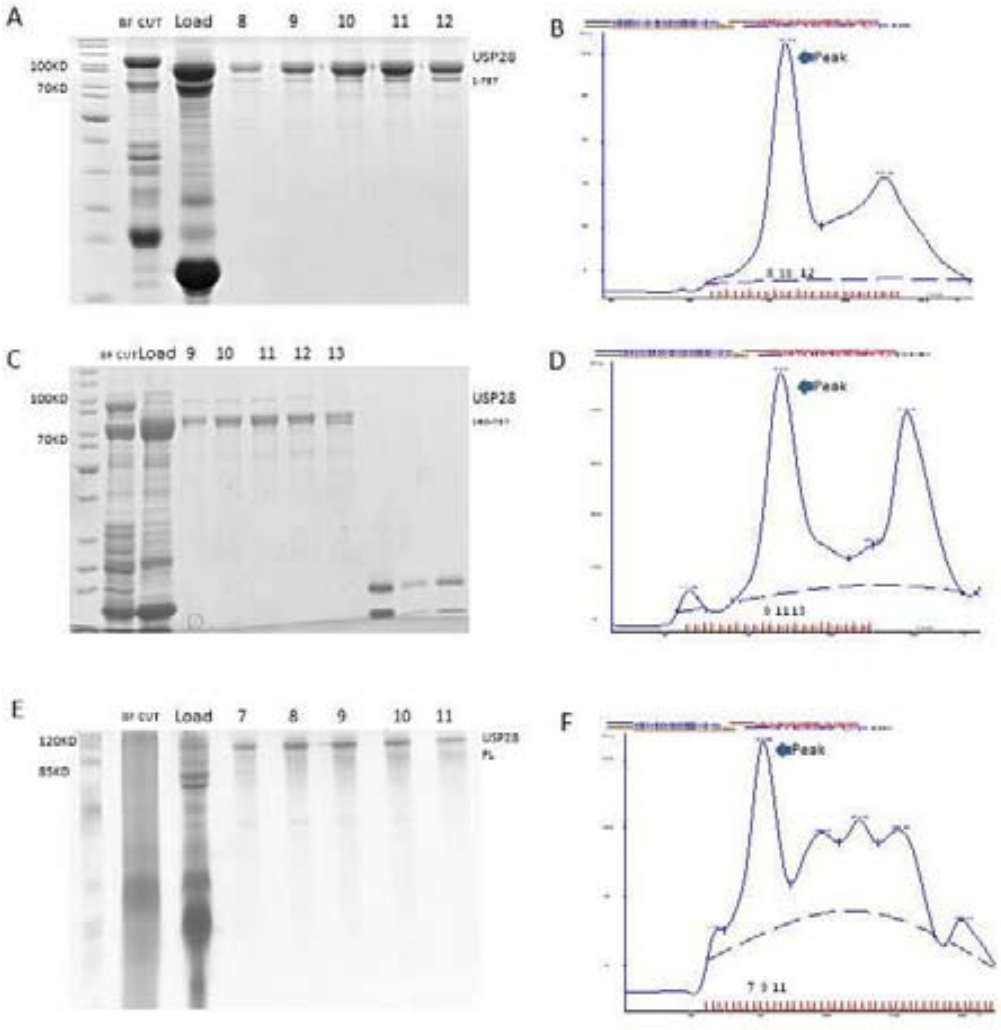


Figure14 –The purification of the new CD+NT of USP28 (USP28 $_{1-757}$), the new CD of USP28 (USP28 $_{160-757}$) and the FL of USP28. A. The SDS-PAGE gel of the USP28$_{1-757}$ gel filtration: bf cut: sample before the SENP cleavage of the Smt3 fusion tag, Load: the load sample of gel filtration; B. The gel filtration chromatography of USP28$_{1-757:}$ the peak containing fragments from 8 to 12 is the USP28$_{1-757}$; C. The SDS-PAGE gel of the USP28$_{160-757}$ gel filtration: bf cut: sample before the SENP cleavage of the Smt3 fusion tag, Load: the load sample of gel filtration; D. The gel filtration chromatography of USP28$_{160-757:}$ the peak containing fragments from 9 to 13 is the USP28$_{160-757}$; E. The SDS-PAGE gel of the USP28-FL gel filtration: bf cut: sample before the SENP cleavage of the Smt3 fusion tag, Load: the load sample of gel filtration; F. The gel filtration chromatography of USP28-FL: the peak containing fragments from 7 to 11 is the USP28-FL.

We finally got all the constructs of USP25 and USP28 with the correct C-Terminal and we could carry out the proteolytic activity analysis again. Compared with others, the USP28 $_{1-673}$ had very low activity in our de-ubiquitinating assays (Figure 15A). As has been described for USP25, an extension of the C-terminal catalytic domain to Tyr757 yielded a recombinant protein with de-ubiquitinating activities comparable to the full-length USP28 and USP25 (figure 15A & 15B). Thus, in both USP25 and USP28 the ~~–c~~onserved" catalytic domain is longer in comparison to other USP family members (Denuc *et al*, 2009).

In our de-ubiquitinating *in vitro* assays with the USP28 constructs we do not observe substantial differences in the proteolytic activities between K48 and K63-linked poly-ubiquitin chains. Interestingly, our *in vitro* assays indicate that removal of the N-terminal region (USP28 $_{160-757}$) does not impair the de-ubiquitinating activity for either the K48 or K63-linked poly-ubiquitin chains substrates, and we

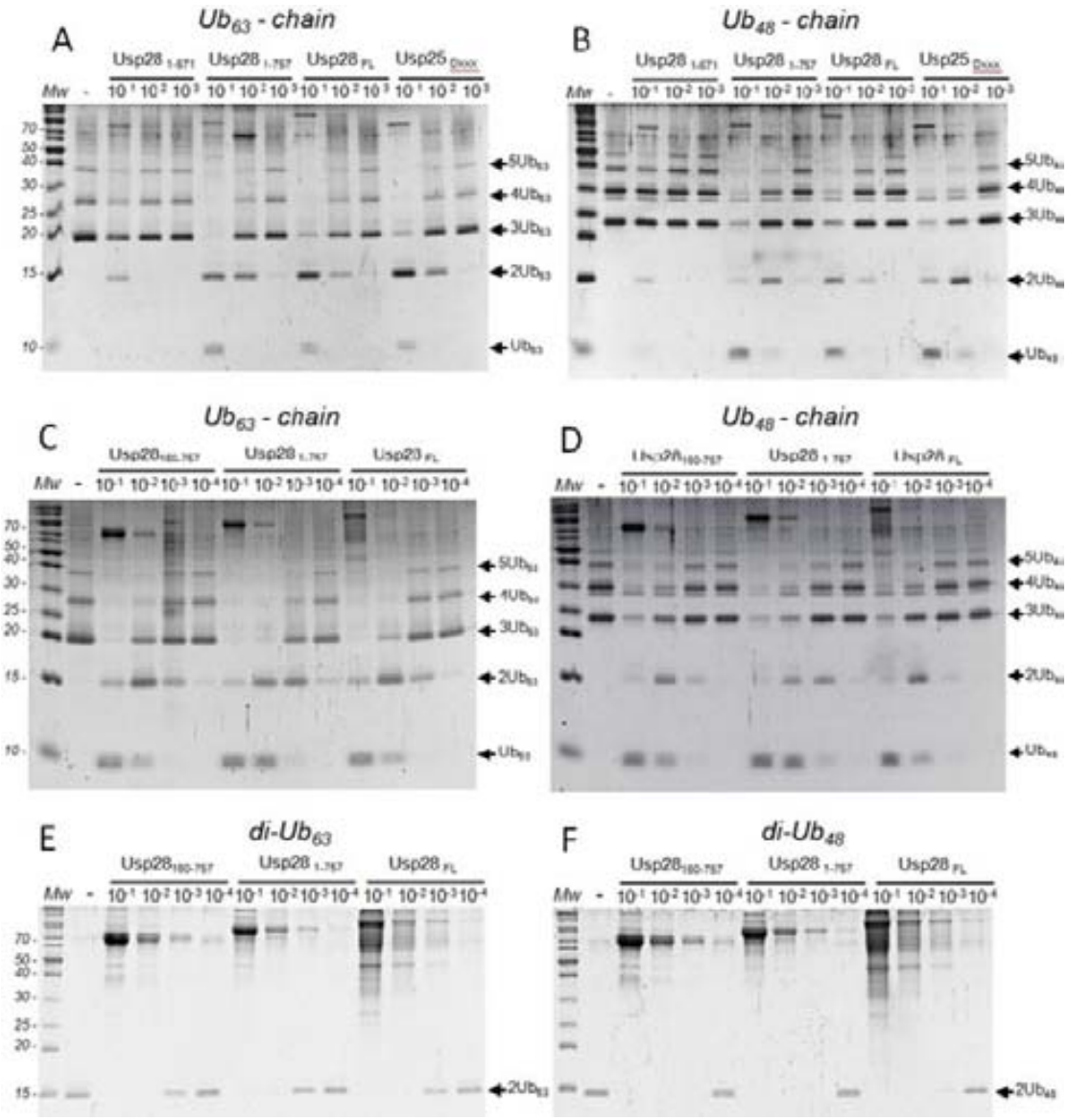


Figure15 - De-ubiquitinating activity of USP28 on K48 and K63 Ubiquitin substrates. A. End point activities on K63-linked Ubiquitin chains of three different dilutions of indicated USP28 constructs after 30 minutes reaction. B. End point activities on K48-linked Ubiquitin chains of three different dilutions of USP28 constructs after 30 minutes reaction. C. End point activities on K63-linked Ubiquitin chains of four different dilutions of indicated USP28 constructs after 30 minutes reaction. D. End point activities on K48-linked Ubiquitin chains of four different dilutions of indicated USP28 constructs after 30 minutes reaction. E. End point activities on K63 di-Ubiquitin substrate of four different dilutions of indicated USP28 constructs after 30 minutes reaction. F. End point activities on K48 di-Ubiquitin substrate of four different dilutions of indicated USP28 constructs after 30 minutes reaction.

cannot detect significant differences in comparison to the activities displayed by the USP28 constructs containing the N-terminal region, namely USP28 $_{1\text{-}757}$ and USP28 FL (Figure 15C &15D). Thus our results indicate that despite the presence of ubiquitin-binding domains at the N-terminal region of USP28, the absence of such region does not affect the de-ubiquitinating activity of USP28, at least in the activity against K48 and K63-linked poly-ubiquitin chains substrates *in vitro*, which is similar conclusion with the USP25 case made by Denuc and coworkers (Denuc *et al*, 2009).

It could be that substrate interaction with the N-terminal domain could lead to potential proteolytic differences only observed when using shorter substrates, such as di-ubiquitin with a single ubiquitin linkage. However, de-ubiquinating analysis using K48 or K63 di-ubiquitin substrates yielded similar results to those using poly-ubiquitin chains (Figure 15E &15F), also indicating that the N-terminal region of USP28 is not a major determinant of activity in our *in vitro* de-ubiquitinating assays. In summary, our results indicate that there is not a particular preference for the most common K48 and K63 di-ubiquitin linkages, and that the presence of the N-terminal region of USP28 that contains several ubiquitin-binding motifs does not affect the de-ubiquitinating activity of the catalytic domain.

1.7 Di-Ubiquitin chain specificity

In order to further investigate the specificity of our USP28 constructs we have tested the de-ubiquitinating ability against all eight possible di-ubiquitin substrates, including the linear, Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63 di-ubiquitin (Figure 17). Time course experiments were run at fixed substrate and USP28 concentrations, 5 μM and 120 nM, respectively. Under this experimental condition only three di-ubiquitin substrates were cleaved by USP28, namely Lys11, Lys48 and Lys63, and in all instances with comparable activities between the three constructs tested, USP28 $_{160\text{-}757}$, USP28 $_{1\text{-}757}$ and USP28 FL. Interestingly, a previous report with several USP proteins, which includes the homologous USP25, showed only minor differences against all eight di-ubiquitin substrates (Faesen *et al*, 2011). In contrast, in the case of USP28 the de-ubiquitinating activity against di-ubiquitin linkages resembles the OTU DUBs family, where members are specific for one or a small subset of di-ubiquitin linkages types (Mevissen *et al*, 2013). It is interesting that in addition to the most common Lys48 and Lys63 linkages, Lys11 shows up in our di-ubiquitin cleavage analysis by USP28 (Figure 16). Lys11 poly-ubiquitin chains have been recently implicated as an alternative degradation signal through the anaphase-promoting complex (APC) during the cell cycle progression in mitosis (Wickliffe *et al*, 2011).

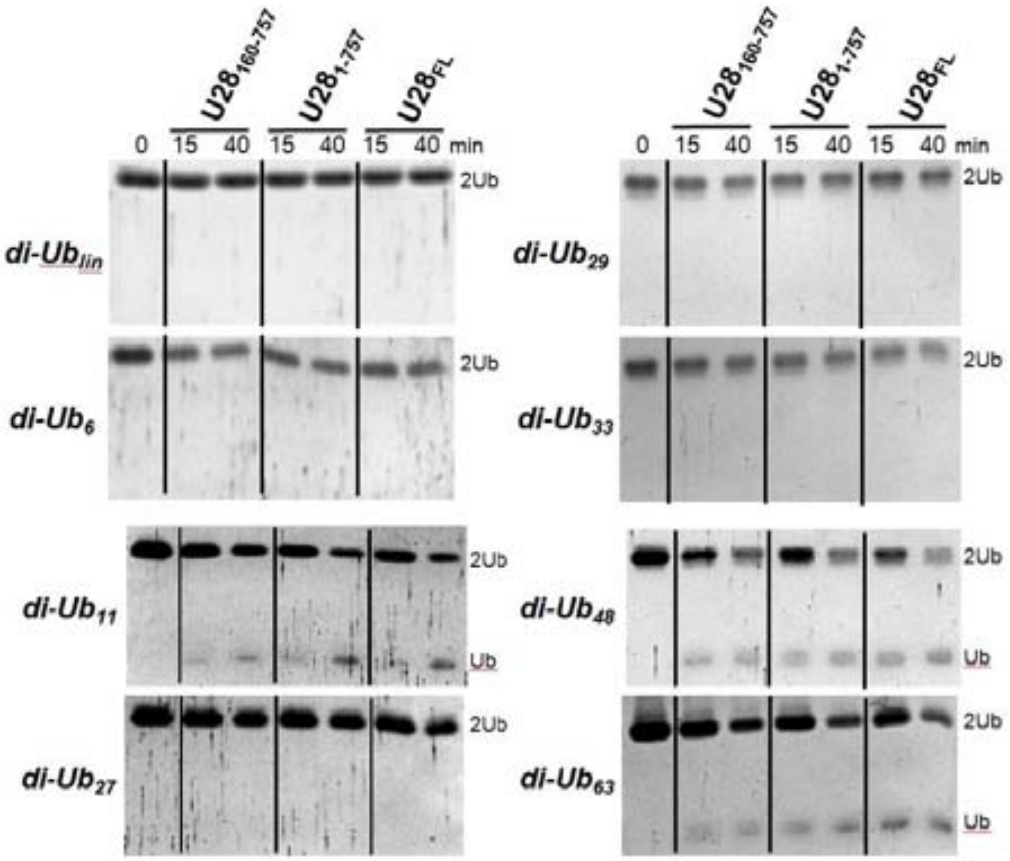


Figure 16 - Di-ubiquitin linkage preference for USP28 constructs. USP28 $_{160\text{-}757}$, USP28 $_{1\text{-}757}$ and USP28 FL constructs were incubated with di-ubiquitin substrates of all linkage types (linear, K6, K11, K27, K29, K33, K48 and K63) for the indicated times and resolved in an SDS-PAGE stained with SYPRO-Ruby. All USP28 constructs were used at a fixed concentration (approx. 150 nM).

1.8 The small scale SUMOylation of the Catalytic Domain (CD) of USP28

Similarly to the N-terminal domain of USP28, we also set up a small scale SUMOylation reaction of the Catalytic Domain (CD) of USP28 using purified E1, E2 and IR1 (E3 SUMO ligase), which resulted in an efficient attachment of SUMO2 to USP28$_{1\text{-}757}$ (Figure 17).

Figure 17 – The SUMOylation of the Catalytic Domain (CD) of USP28. The first lane is the protein ladder. The second lane is the time 0‘ without adding the ATP. Then the samples are taken at 30‘, 60‘ and 120‘.

1.9 USP28 SUMOylation primary site identification by the MS and mutagenic SUMOylation analysis.

In the previous experiment, we have successfully performed the SUMOylation of the Catalytic Domain (CD) of USP28. Similarly to the N-terminal of USP28, we cut the SUMOylated USP28-CD band from the gel and sent it to the Mass Spectrometry facility of the IRB in Barcelona to check whether the primary SUMOylation sites were contained in the N-terminal region. MS analysis of an excised band of an *in vitro* SUMO conjugation reaction of a construct including the N-terminal region and the conserved catalytic domain of USP28 (USP28 $_{1\text{-}757}$), also confirmed Lys99 as the major SUMOylation site (Table 4). The analysis also revealed the presence of other low level SUMOylation sites, including Lys115, Lys385, Lys511 and Lys513.

Residue	Non-modified (PSM#)	SUMOylated (PSM#)	Motif ψKXE
K64	47	1	VKEP
K85	82	1	NKEV
K99	82	23	NKDaD
K115	101	4	PKIQ
K210	0	2	EKRN
K305	75	1	GKPF
K385	62	1	NKLE
K446	183	1	LKYV

Table 4 - The Mass Spectrometry analysis of SUMOylated USP28-NT. Sites identified in SUMOylated fragments of the USP28 $_{1\text{-}673}$ protein. Number of peptide spectrum matches (PSMs) for the non-SUMOylated (left column) or SUMOylated (right column). Motif compared to linear consensus motif is shown for each (far right colum).

To confirm the MS results, *in vitro* SUMO conjugation reactions were also conducted with two different SUMO E3 ligases, IR1 and Nse2, and with two point mutants of USP28, K99R and K115R. Again, Lys99 emerges as the major SUMOylation site, although SUMO conjugation can also occur in a lesser degree in Lys115 (Figure 18A and 18B). Double point mutant K99R and K115R practically eliminates SUMO conjugation on USP28 in the two separate reactions using different E3 SUMO ligases. Interestingly, in our SUMO conjugation assays, the absence of the

N-terminal region ($USP28_{160-757}$) also reduces the formation of SUMO conjugates, indicating that the N-terminal region of USP28, and in particular Lys99, is the major site for SUMO conjugation in our *in vitro* assays (Figure 18A and 18B, last lane).

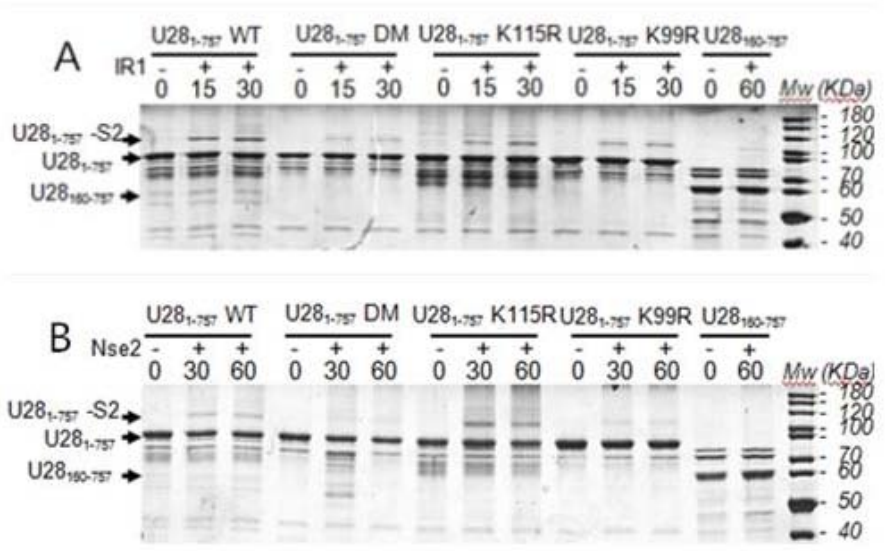


Figure 18 - SUMO conjugation reaction with USP28-CD-Mutants. A. Time course SUMO conjugation reaction with IR1 E3 ligase using different point mutants of the catalytic domain construct of USPS28 ($USP28_{1-757}$). B. Similar reaction as in A but using Nse2 as SUMO E3 ligase. Reactions were run at 37 C and stopped with SDS-PAGE loading buffer at marked times. $USP28_{1-159}$ K64R/K115R (DM: Double Mutant). $USP28_{1-159}$ K64R/K115R/K99R (TM: Triple Mutant). $USP28_{1-757}$ K99R/K115R (DM: Double Mutant).

1.10 Characterization of SUMO modification on USP28 proteolytic activity.

We next examined the effects of SUMO modification on USP28 proteolytic activity using a deubiquitination assay. We first set up a large-scale SUMO conjugation reaction using IR1 as a SUMO E3 ligase, followed by an ion exchange chromatography to separate non-modified USP28 from the SUMO conjugated to USP28 (data not shown). This step is essential to reduce, as much as possible, any contamination of USP28 from the SUMO-USP28 preparation. As shown in the penultimate lane in Figure 19B, the USP28 band is hardly observed in a Ruby-SYPRO stained gel. However, we cannot discard some minor level of contamination of free USP28 in the SUMO conjugated preparation.

In order to assure equal amounts of USP28 protease in our comparative analysis, a serial dilution of USP28-SUMO conjugate was prepared with and without the presence of the SUMO protease SENP2. After proper cleavage of SUMO from USP28, the de-ubiquitinating activity was checked using the K48 and K63 poly-ubiquitin chains and di-ubiquitin substrates. Comparison in the processing of poly-ubiquitin chains substrate before and after SENP2 treatment clearly indicates a reduced activity in the USP28-SUMO adducts (Figure 19A & 19B). Only at high protease concentrations can the SUMO-modified USP28 efficiently cleave the K48 or K63 poly-ubiquitin chains. This residual activity at high concentration could be a consequence of contamination of non-modified USP28 in the SUMO-USP28 preparation or differences in the total SUMOylation levels of multiple lysines. We again did not observe significant differences in the proteolytic activity between the

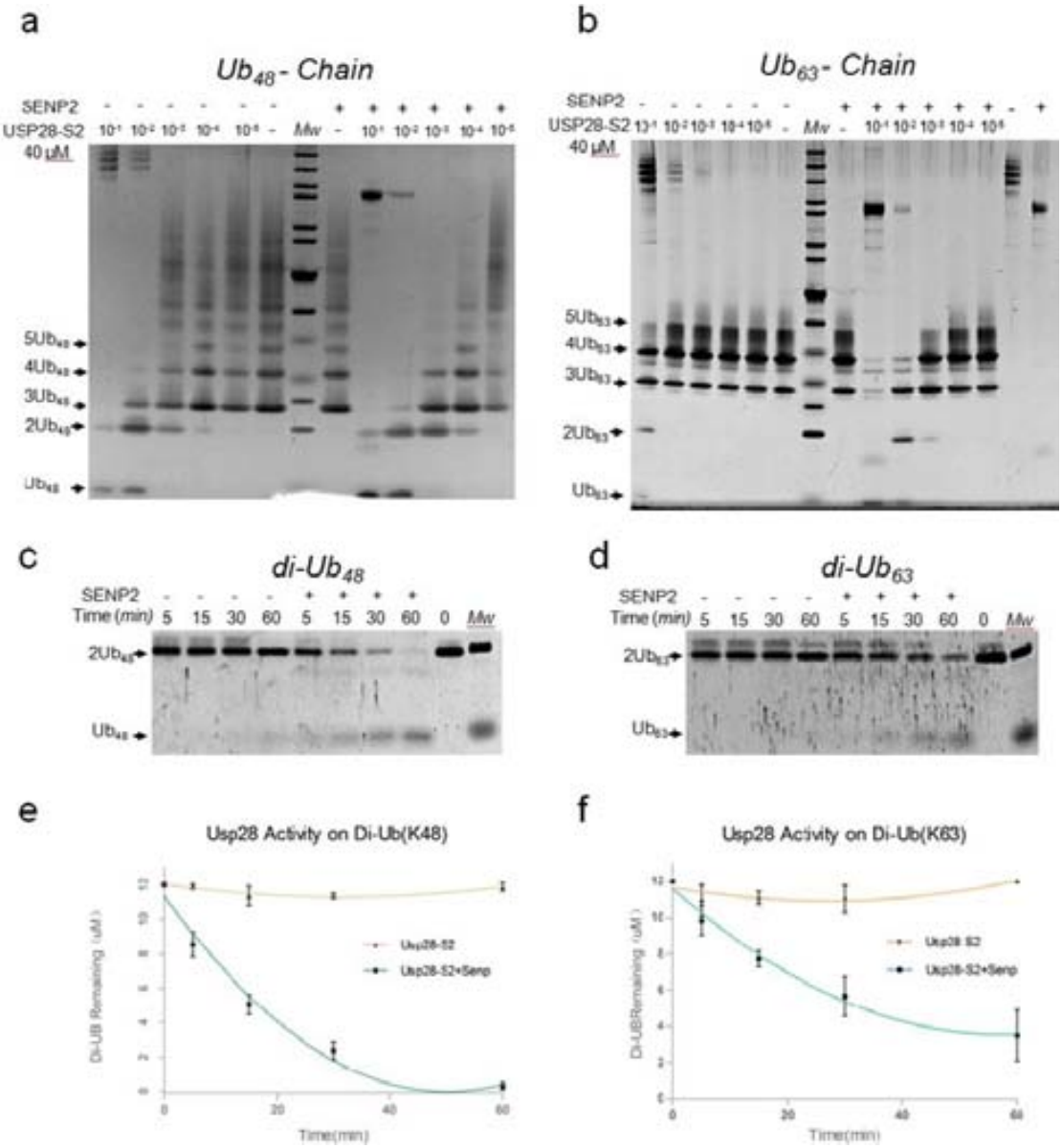


Figure 19 – Inhibition of the de-ubiquitinating activity of USP28 by SUMO conjugation. A. End point de-ubiquitinating activity comparison on K48-linked Ubiquitin chains using different dilutions of SUMO conjugated to the USP28$_{1-757}$ construct, before and after SENP2 treatment. B. End point de-ubiquitinating activity comparison on K63-linked Ubiquitin chains using different dilutions of SUMO conjugated to the USP28 $_{1-757}$ construct, before and after SENP2 treatment. C. Time course reaction of de-ubiquitinating activity on K48 di-Ubiquitin substrate using the SUMO-conjugated USP28 $_{1-757}$ construct, before and after SENP2 treatment. D. Time course reaction of de-ubiquitinating activity on K63 di-Ubiquitin substrate using the SUMO-conjugated USP28 $_{1-757}$ construct, before and after SENP2 treatment. E. Graphic representation of the de-ubiquitination reaction shown in C. F. Graphic representation of the de-ubiquitination reaction shown in D. Reactions was running in triplicate.

two poly-ubiquitin chains utilized, K48 or K63, indicating than in both cases, SUMO modification of the N-terminal region inhibits USP28 activity in our *in vitro* de-ubiquitinating assays.

These results with poly-ubiquitin chains can also be recapitulated using K48 and K63-linked di-ubiquitin substrates. In order to have a more quantitative assay, we performed a time-course analysis up to 60 minutes using a fixed concentration of USP28-SUMO, before and after treatment with SENP2 protease(Figure 19C & 19D). Similar to the results with poly-ubiquitin chains, we observed a diminished activity when USP28 was SUMO conjugated, compared to the activity after treatment with SENP2. Interestingly we noted that after a digestion of 60 minutes, the K48 di-ubiquitin substrate is completely digested, while the proteolysis of the K63-linkage occurs at a slower rate (Figure 19E & 19F). Although differences in our *in vitro* assays are subtle, they might indicate a better interaction between USP28 and the K48-linkage, at least for the di-ubiquitin substrates.

In USP25, it was described that a covalent and a non-covalent interaction of SUMO with the N-terminal region of the protease, resulted in an inhibition of de-ubiquitinating activity of USP25 (Meulmeester *et al*, 2008). In order to check a potential non-covalent SUMO inhibition mechanism on USP28, we have conducted a competitive *in vitro* de-ubiquitinating activity assay in the presence of SUMO2 (figure 20). For this assay, increasing amounts of SUMO2 were added to a fixed concentration of the two truncated constructs of USP28, the N-terminal domain plus catalytic domain (USP28 $_{1\text{-}757}$), and only the catalytic domain (USP28 $_{160\text{-}757}$). The

processing of the K48 and K63-linked poly-ubiquitin chain substrates do not vary significantly after the addition of increasing amounts of SUMO2 (Figure 20A and 20B). We do not observe any inhibition of the USP28 de-ubiquitinating activity in the presence of the N-terminal domain, even when SUMO2 concentrations are several orders of magnitude higher. These results indicate that the covalent linkage formed between SUMO and the N-terminal region of USP28 (primarily through Lys99) is necessary for the inhibition of USP28 catalytic activity, and that in our *in vitro* assays a non-covalent interaction of SUMO with the N-terminal region of USP28 does not affect the activity of the protease.

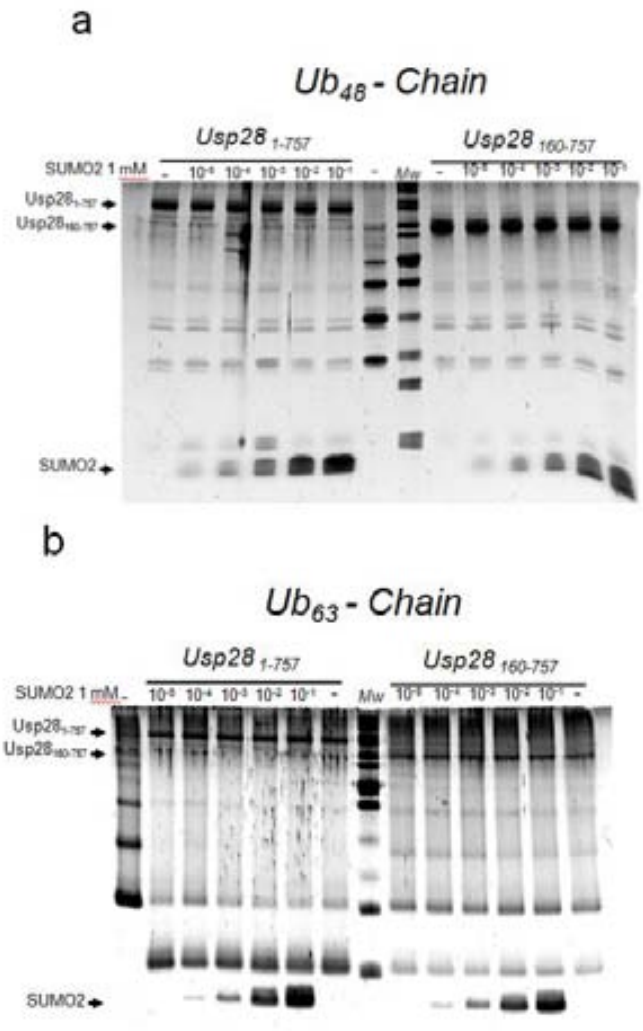


Figure 20 – Non-covalent competitive analysis of SUMO2 on the de-ubiquitinating activity of USP28. A. Analysis of the de-ubiquitinating activity of the USP28 $_{1-757}$ and USP28 $_{160-757}$ constructs on K48-linked Ubiquitin chains in the presence of increasing concentrations of SUMO2. B. Analysis of the de-ubiquitinating activity of the USP28 $_{1-757}$ and USP28 $_{160-757}$ constructs on K63-linked Ubiquitin chains in the presence of increasing concentrations of SUMO2.

1.11 Crystallization of USP25-CD and USP28-CD constructs.

We next set up trays of the USP25-CD and USP28-CD to try to form crystals using the crystallographic high-throughput robot for the lab. Both individual proteases were sprayed into 96-well sitting drop plates, mixed with commercial screen conditions and stored at 18°C. We only could find crystals of USP25$_{1-739}$ showing up in a few days in some of the conditions (Table 5) (Figure 21), whereas we could not observed any crystals for USP28. We then optimized the conditions in 24-well-plates made by mixing 0.5-1μL of USP28$_{1-159}$ with corresponding buffer and adding to sitting drop plates and stored at 18°C. After incubation several days, bigger crystals showed up and were fished up and stored in liquid nitrogen. The X-ray diffraction analysis was carried out at the Alba synchrotron. Unfortunately, although this time we could collect a complete diffraction data set, the diffraction signal was only around 7Å which is not good enough to solve the structure at atomic level (Diffraction data in Table 6).

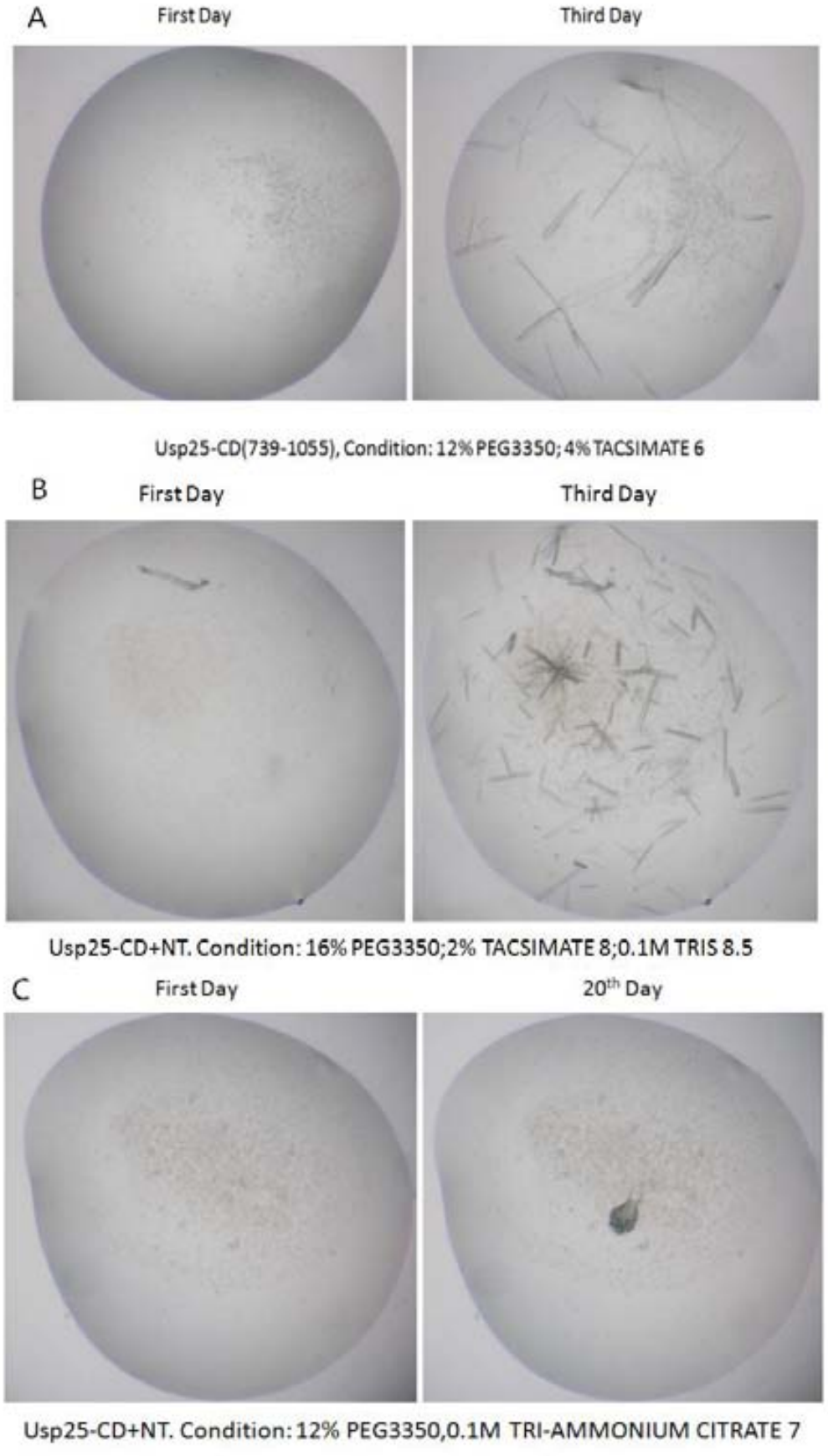
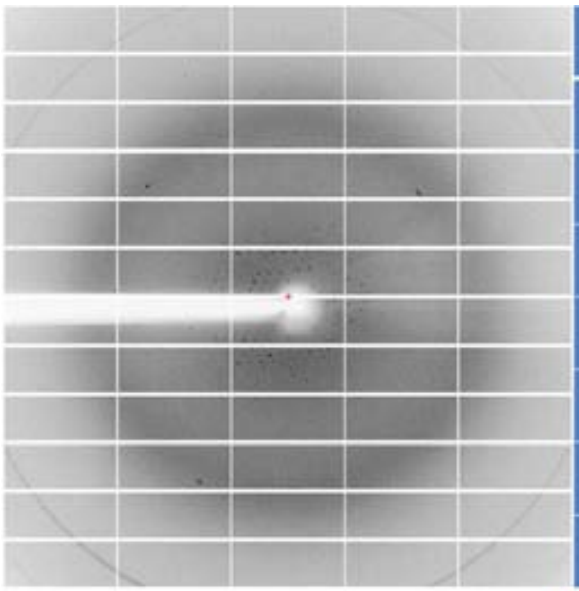


Figure 21 – Crystals of USP25-CD.

Screens	Conditions	Showing up
Proplex	0.2M NaCl, 12% PEG8000, 0.1M HEPEs PH 7.5	2nd Day
Proplex	0.2M NH_4SO_4, 12% PEG8000, 0.1M Tris PH 7.5	2nd Day
Morpheus	0.06M Divalents, 0.1M Sys1 PH 6.5, 30% GOL_P4K	2nd Day
Morpheus	0.06M Divalents, 0.1M Sys2 PH 7.5, 30% 30% GOL_P8K	2nd Day
Morpheus	0.09M NPS, 0.1M Sys3 PH 8.5, 30% EDO_P4K	2nd Day
Morpheus	0.1M Amino acids, 0.1M Sys3 PH 8.5, 30% GOL_P4K	2nd Day
PEG/ION	0.1M Na Malonate PH 6 or 7,12% PEG3350	2nd Day
PEG/ION	0.1M Malic Acid PH 7,12% PEG3350	2nd Day
PEG/ION	5% Tacsimate PH6,7 and 8, 12% PEG3350	2nd Day
PEG/ION	0.1M $(NH_4)_3$Citrate PH7, 12% PEG3350	2nd Day

PEG/ION	0.1M $(NH_4)_3$Tartrate PH7, 12% PEG3350	2nd Day
JCSG PLUS	0.2M $MgCl_2$, 12% PEG8000, 0.1M Tris PH 7.5	2nd Day
JCSG PLUS	0.1M Potassium thiocyante, 30% PEG2000MME	2nd Day
JCSG PLUS	0.1M Potassium Bromide, 30% PEG2000MME	2nd Day
Index	0.2M KCl, 0.05M HEPEs PH7.5, pentacrythritol propoxylate(5/4 PO/OH)	2nd Day
Index	0.1M Succinic acid PH 7, 15% PEG3350	2nd Day

Table 5 - The screens and conditions of $USP28_{1-159}$ crystals.



Data Collection	USP25-CD+NT
Space group	I422
a, b, c (Å)	139.4, 139.9, 191.9
α, β, γ (°)	90, 90, 90
Resolution (Å)*	50 – 7 (6.88– 7.15)
R_{merge}*	0.152
I/σI	16.29
Completeness (%)	99.5 (99.2)

Table 6 - USP25-CD diffraction pattern and data set characterization.

2. The Results of SENP part

2.1 Structural characterization and purification of SENP3 and SENP5 protease.

As mentioned in the introduction, several studies have been already published on the characterization of the SENP1/2 sub-family, however for the SENP3/5 members the information is still scarce. Both SENP3 and SENP5 are proposed to be localized at the nucleolus, where the SUMO2/3 is also localized. A few studies reported SENP3 and SENP5 to have specificity for SUMO-2 and SUMO-3 isoforms and much less for SUMO-1 (Gong *et al* 2006). Either the crystal structures of the long N-terminal extension (NT) or of the Catalytic Domain (CD) of SENP3/5 are still unknown.

Based on structural alignments with other members of the SENP protease family (Figure 22A), and on a previously published reports on the homologous SENP2 and SENP1 (Figure 22B and C, Reverter *et al*., 2006 and Hay, 2006), the SENP5 full-length protein (755 residues) can be divided to two domains: the non-homologous N-terminal domain (around 300 residues long) and the catalytic ~~"e~~conserved" SENP domain (around 500 – 755 residues long).

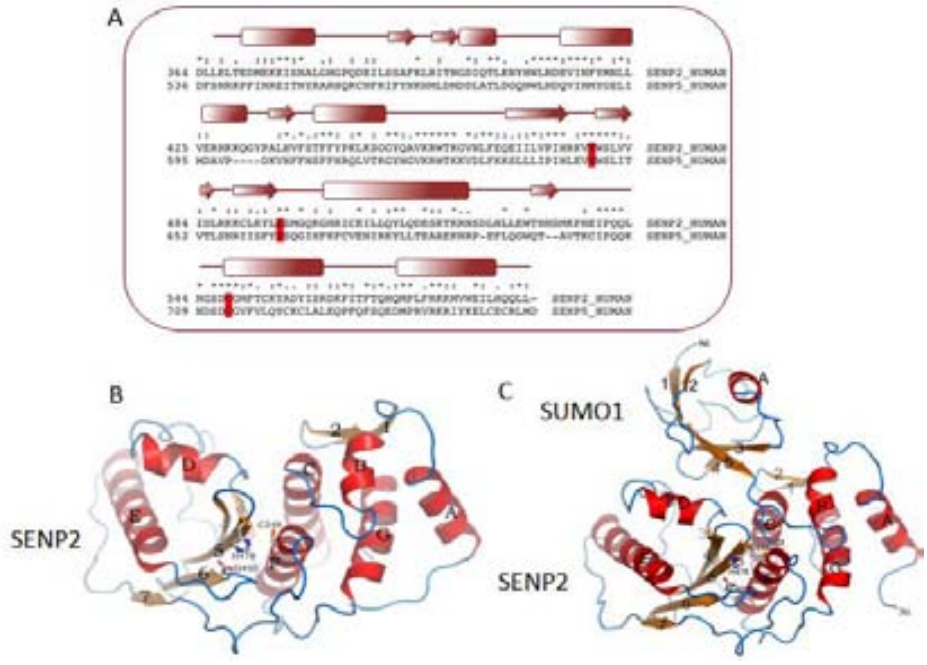


Figure 22 *- A. Sequence alignment of SENP2 and SENP5 catalytic domains*. Secondary structural elements are based on SENP2 structure (Reverter D, 2006) and shown below the sequences. *B. The SENP2 Structure. C. The complex between SENP2 with SUMO2 precursor.*

A possible reason for the less and limited information about SENP3 and SENP5 is partly because they are hard to be purified in *E. coli* cells. Thus a major goal in my work is to be able to purify members of this SENP sub-family in high yields in order to solve the crystal structure of either SENP3 or SENP5 by X-ray crystallography.

Based on the structural alignments with of SENP3 and SENP5 with other members of the SENP family, we have designed Catalytic Domain constructs of different lengths of SENP3 and SENP5 in order to obtain soluble proteins. The analyzed constructs included: SENP3-NΔ92, SENP3-NΔ257, SENP3-NΔ290, SENP3-NΔ354, SENP3-NΔ377 and SENP3-NΔ390; SENP5-NΔ440, SENP5-NΔ475, SENP5-NΔ535, SENP5-NΔ559 and SENP5-NΔ568 (Figure 23). Among all these constructs, the plasmids containing the constructs SENP3-NΔ92, SENP3-NΔ290, SENP3-NΔ354, SENP5-NΔ475, SENP5-NΔ535 and SENP3 and SENP5-FL were already made at the Memorial Sloan-Kettering Cancer Center in New York by David Reverter. All the other constructs were designed cloned and analyzed in the present work with the general purification method described in the Methods section of this thesis.

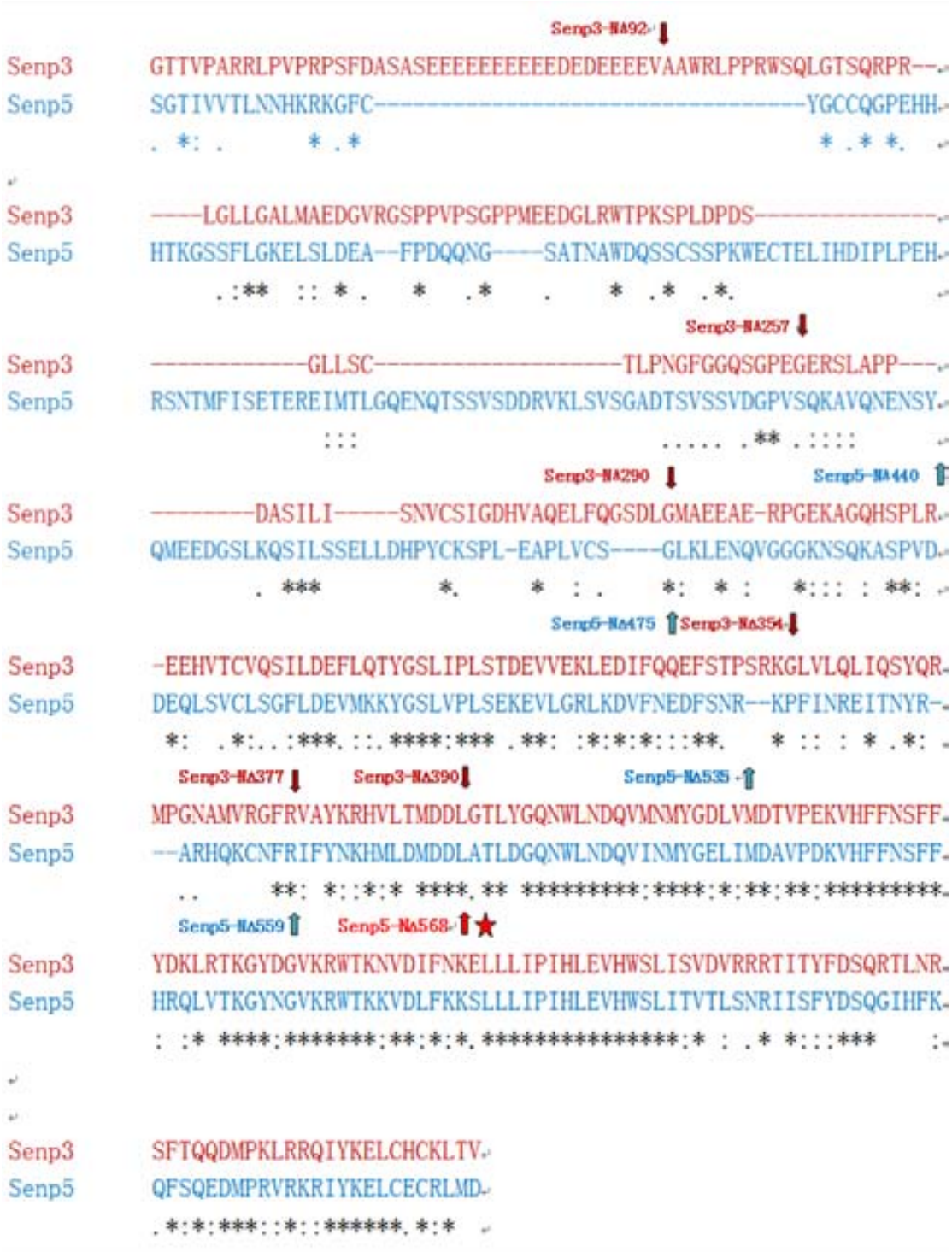


Figure 23 – The full length sequence alignment and constructs of SENP3 and SENP5. The small stars and dots indicate the conserved residues. The arrow point to the last residue of constructs what we made and used in this thesis. The red star below the SENP5-NΔ568 indicates the only construct we can express soluble in *E. coli*.

In figure 23, the comparison of the primary sequences of SENP3 and SENP5 is shown, indicating the length of the constructs that have been check in this work. Unfortunately, as mention before, in contrast to SENP1 or SENP2 the Catalytic Domains of SENP3 and SENP5 are really hard to be expressed in *E. coli* cells. The different constructs analyzed, including different fusion tags, are listed in the table bellow (Table 7).

Vectors/Fusion Tags	Expression Strains	CD OF SENP3 and SENP5	Expression
PET28b, His	BL21(DE3)	SENP3-NΔ92	NO
PET28b, His	BL21(DE3)	SENP3-NΔ290	NO
PET28b, His	BL21(DE3)	SENP3-NΔ354	NO
PET28b, His	BL21(DE3)	SENP3-NΔ377	NO
PET28b, His	BL21(DE3)	SENP3-NΔ390	NO
pGex4T-1, GST	BL21(DE3)	SENP3-NΔ390	NO
pGex4T-1, GST	BL21(DE3)	SENP3-FL	NO
pNusA, NusA	BL21(DE3)	SENP3-FL	NO
pMBP, MBP	BL21(DE3)	SENP3-FL	NO
PET28b, His	BL21(DE3)	SENP5-NΔ440	NO
PET28b, His	BL21(DE3)	SENP5-NΔ475	NO
PET28b, His	BL21(DE3)	SENP5-NΔ535	NO
PET28b, His	BL21(DE3)	SENP5-NΔ559	NO

pGex4T-1, GST	BL21(DE3)	SENP5-440	NO
pGex4T-1, GST	BL21(DE3)	SENP5-FL	NO
pNusA, NusA	BL21(DE3)	SENP5-FL	NO
pMBP, MBP	BL21(DE3)	SENP5-FL	NO
PET28b, His	BL21(DE3)	SENP5-NΔ568	NO
PET28b, His	Rosetta 2, PlysS	SENP5-NΔ568	NO
pGex4T-1, GST	BL21(DE3)	SENP5-NΔ568	NO
PET28b, His	SoluBL21(DE3)	SENP5-NΔ568	NO
PET28b, His	Rosetta 2	SENP5-NΔ568	YES

Table 7 - Expression of different Catalytic Domains of SENP3 and SENP5 in different vectors.

Among all these constructs analyzed, only the shortest construct of SENP5 (SENP5-NΔ568) could be produced efficiency in a PET28b-His vector in the *E. coli* strain *Rosetta 2*.

Due to the easy precipitation of the SENP5-NΔ568 construct during the scale-up purification after the thrombin-cleavage tag in the nickel-nitrilotriacetic acid-agarose resin (NTA-Ni^{2+}, Qiagen), we used a different purification method. First we directly let the NTA-Ni^{2+} eluate pass through a Gel Filtration column (Superdex 75; GE Healthcare) and collect the fractions containing the His-tag fusion protein His-SENP5-NΔ568. Next we digest it with Thrombin at 1:1000 ratio for 2h at 37 °C to remove the His-Tag and then run a second Gel Filtration (Superdex 75; GE

Healthcare) to get the single SENP5-NΔ568 protein, which corresponds about 20 KDa in the SDS-PAGE gel, as expected form the theoretical molecular weight (Figure 24).

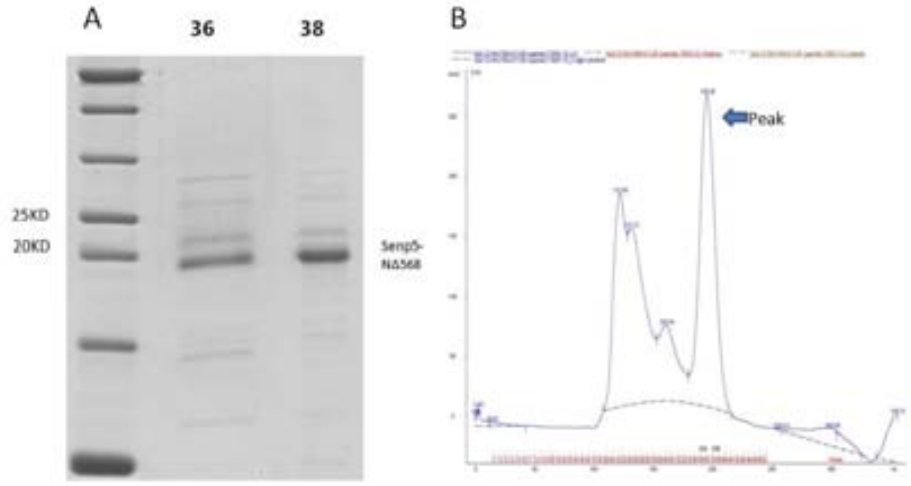


Figure 24 – The purification of SENP5-NΔ568. A. The SDS-PAGE gel of the SENP5-NΔ568 gel filtration; B. The gel filtration chromatography of SENP5-NΔ568: the peak containing fragments from 36 to 38 is the SENP5-NΔ568 protein;

2.2 Functional characterization of the SENP5 catalytic domain.

As we know, the SUMO protease SENP family contains both functions of processing and deconjugation. Like ubiquitin, SUMO proteins are expressed as precursor proteins that carry a C-terminal extension of variable length (2–11 amino acids) found after a conserved di-glycine motif (Figure 25). To be able to function as a modifier of target proteins, the C-terminal di-glycine motif of the SUMO proteins has to be exposed by the action of SENPs, this reaction is named ~~"p~~rocessing reaction" (figure 25).

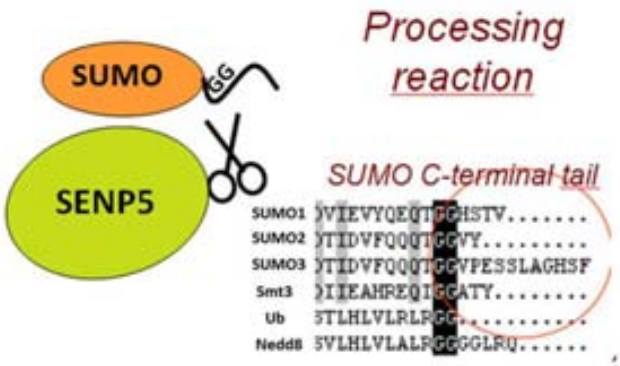


Figure 25 – The SUMO processing ability. Alignment indicates the C-Terminal tail of the sumo binding region.

Thus we first use the SUMO-precursors (full-length pSUMO1, pSUMO2 and pSUMO3) as the substrates to verify this SENP5-CD (SENP5-NΔ568) "processing" activity.

The pSUMO1, pSUMO2 and pSUMO3 plasmids were constructed at the Memorial Sloan-Kettering Cancer Center in New York City by David Reverter. They were purified with the general procedure detailed in the Methods. The SUMO precursors were diluted 10 times to a final 0.1 mg/ml and mixed with different concentrations (150nM to 0.15nM) of SENP5-NΔ568 at 37 °C in a buffer containing 25 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.1% Tween 20, and 1 mM dithiothreitol. Processing reactions were stopped after 25 min with SDS loading buffer and analyzed by gel electrophoresis (SDS-PAGE). In figure 26 we can observe that although SENP5-NΔ568 has the ability to cut all SUMO precursors, the processing ability on SUMO2 is better compared to SUMO1 and SUMO3 (Figure 26), indicating a possible role of the C-terminal tail of the precursors (keep in mind that SUMO2 and SUMO3 contain are identical except on their C-terminal tail).

Figure 26 – The SUMO precursors processing by SENP5-NΔ568. *The SENP5 concentration is from 150nM to 0.15 nM.*

In order to check the role of the C-terminal tail, we next carried out the same "processing" experiment on the pSUMO1GG, pSUMO2GG and pSUMO3GG, which are the same SUMO precursors with an insertion of two additional glycine residues after the Gly-Gly motif (making a total of 4 glycines). These constructs minimize the effect of the different C-terminal tails and are supposed to be easily cut by SENPs compared with the wild type. In this case we can observe that these substrates are really easily cut by the SENP5-NΔ568 and again indicated that the Catalytic Domain of SENP5 has a better processing ability on SUMO2 compared with SUMO1 and SUMO3 (Figure 27).

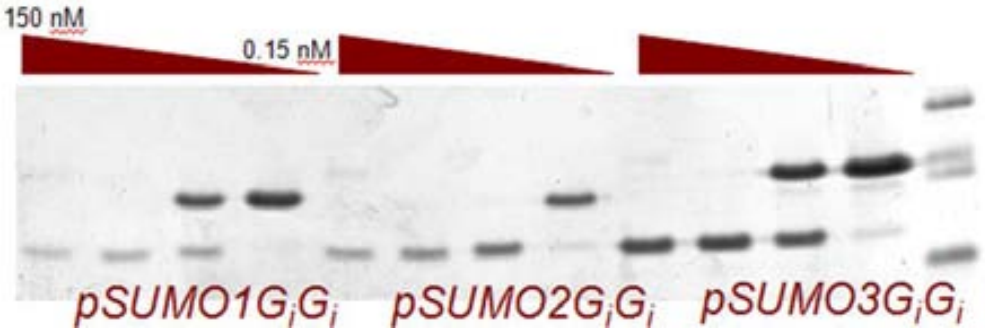


Figure 27 – The SUMO precursors with GG tail processing by SENP5-NΔ568. The SENP5 concentration is from 150nM to 0.15 nM.

Besides the ~~–p~~rocessing" activity, we next analyzed the ~~–d~~econjugation" activity of SENP5-NΔ568, which refers to the removal of SUMO from the SUMOylated substrates (Figure 30). Here, in order to check the deconjugation activity of SENP5-NΔ568, we first set up a SUMOylation reaction on the putative SUMOylation substrate RanGAP1. SUMOylation reaction on RanGAP1 was performed as explained in the Methods section of this thesis. RanGAP1-SUMO2 conjugate can be purified in large amounts by an Anion Exchange column (Resource Q; GE Healthcare). This substrate, RanGAP1-SUMO2, will be used for the deconjugation activity assays of SENP5-NΔ568 and also to form the complex with the inactive catalytic mutant of SENP5-NΔ568 to grow protein crystals and solve the structure (shown later) (Figure 28).

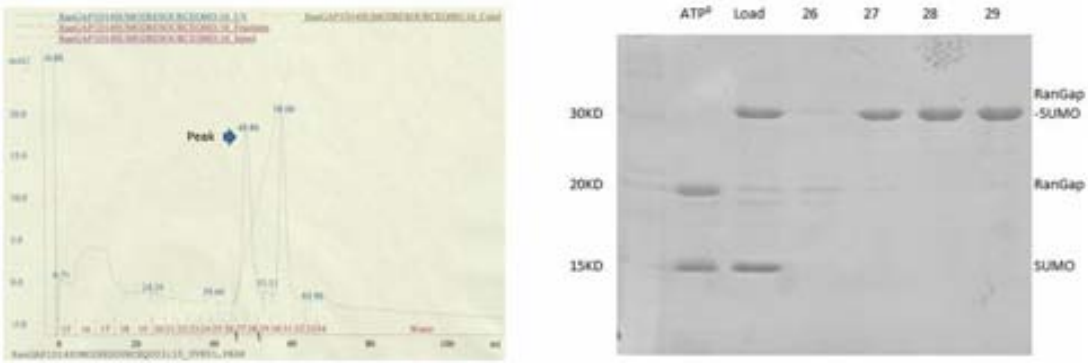


Figure 28 – The SUMOylation and purification of RanGAP1-SUMO. Reaction is hold for 2 hours and almost all the RanGAP1 are sumoylated. The first peak (27-29) is the sumoylated RanGAP1 and the second peak is the rest sumo.

We also set up another SUMOylation reaction to form another SUMO substrate, di-SUMO2. This reaction can be completed by using two different SUMO constructs, Δ14SUMO2 and SUMO2ΔGG, and can be performed as explained in the Methods

section. Finally we are able to purify a large amount of the di-SUMO2 substrate by an Anion Exchange column (Resource Q; GE Healthcare) (Figure 29).

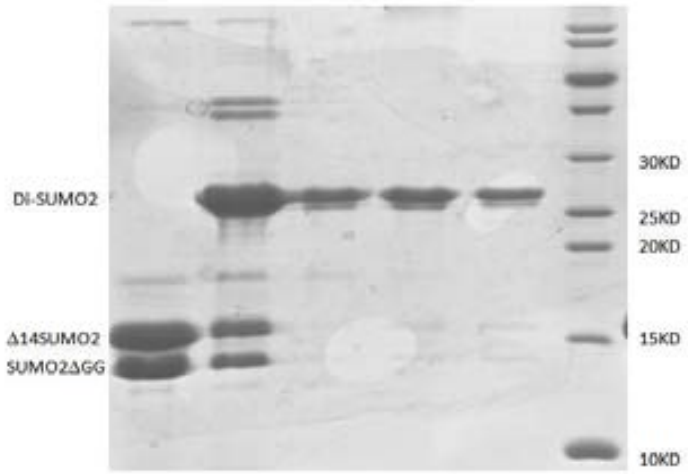


Figure 29 – The SUMOylation reaction and purification of di-SUMO2. The first lane is the ATP time 0; the second lane is the reaction after 2 hours; The 3rd, 4th and 5th lanes are the peak after the Resource Q purification. Half Δ14SUMO2 are sumoylated in 2 hours.

The deconjugation activity of the SENP5-NΔ568 construct has been checked using these two substrates, RanGAP1-SUMO1/2 and di-SUMO2 (Figure 30B). The de-SUMOylation assays were set up as described in the Methods section. First we can observe that SENP5-NΔ568 protease is quite active to cut these two types of SUMOylated substrates, in comparison with the processing reaction. Even when the protease is diluted 10000 times, it can still cut the SUMOylated substrates (Figure 30), this result resembles the activities of SENP1 and SENP2 proteases.

It has reported that SENP5 has specificity for SUMO-2/3 isoforms but less for SUMO-1 (Gong *et al* 2006). In our *in vitro* experiment, we can also observe that when the substrates are RanGAP1-SUMO1 and RanGAP1-SUMO2, SENP5-NΔ568 can cut RanGAP1-SUMO2/3 more efficiently than RanGAP1-SUMO1 (Figure 30C and 30D),

probably indicating a preference of SENP5 for SUMO2/3 isoforms, as reported in *in vivo* experiments (Gong *et al* 2006).

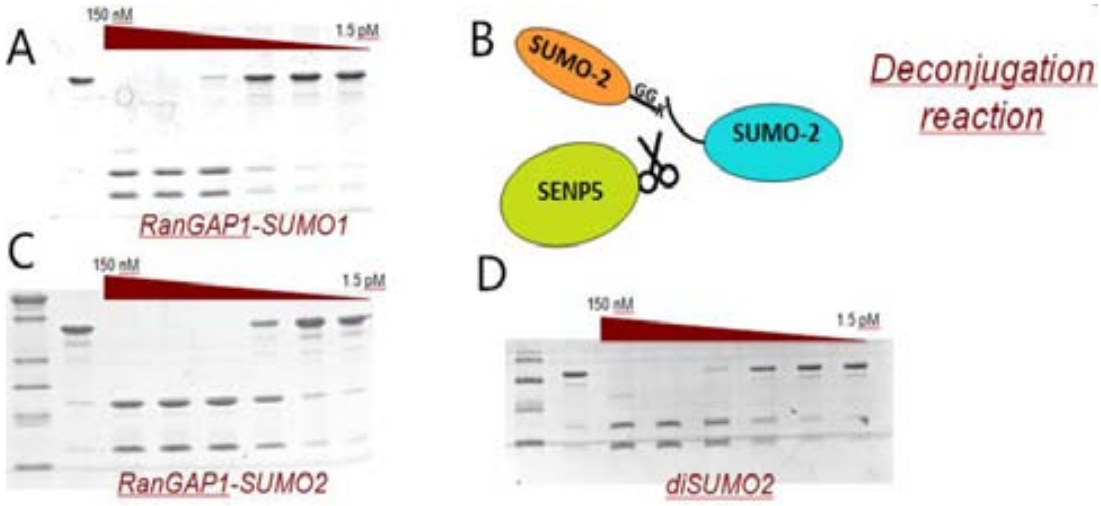


Figure 30 – The deconjugation activity test of SENP5-NΔ568. The SENP5 concentrations range comprises from 150nM to 1.5 pM in all reactions.

2.3 Production and characterization of the inactive mutant of the SENP5-CD-C712S and complex formation between SENP5-CD-C712S with SUMO2 precursor and RanGAP1-SUMO2.

In order to study the interaction between SENP5 and SUMO2 by the X-ray crystallography, first we constructed the inactive mutant of the Catalytic Domain of SENP5-NΔ568 (SENP5-CD-C712S), in which the catalytic active site Cys712 is mutated to serine. Next we first made the proteolytic analysis on RanGAP1-SUMO substrate and we could confirm that this C712S SENP5 mutant construct is really inactive (Figure 31A). To form the complex we mix the SENP5-CD-C712S and the SUMO2 precursor at a 1:2 ratio in a buffer containing 5mM NaCl, 20mM Tris, pH 8.0

and 1mM β-mercaptoethanol. We next pass this mixture through a Gel filtration column (Superdex 200, GE Healthcare) and we confirmed the interaction between these two proteins because they seem to elute together in one single peak (Figure31). A SDS-PAGE gel indicates that the two proteins are eluting together in a stoichiometric proportion almost 1:1 (Figure31B).

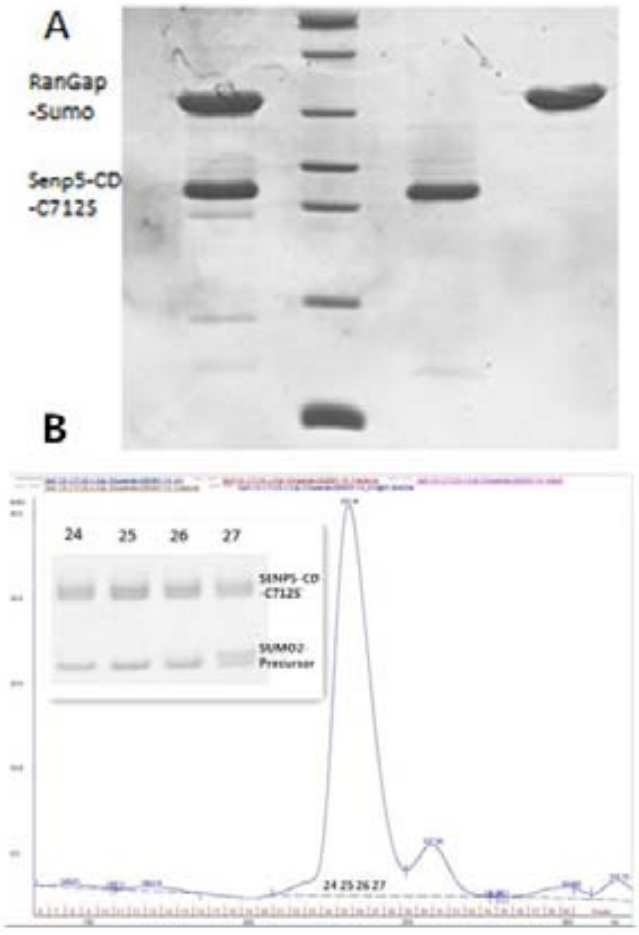


Figure 31 – Inactive SENP5-CD constructing and Complex formation. A. Mix the SENP5-CD-C712S with RanGAP1-SUMO as 1:1 ratio at 37 C in 4 hours to verify that this SENP5-CD is inactive. B. The gel filtration (Superdex 200) chromatography of SENP5-CD-C712S and SUMO2 precursor mixture and the peak from gel filtration running in the SDS-PAGE gel shows that they elute together which means the complex has been formed.

Additionally, to form the second deconjugation complex, we also mix SENP5-CD-C712S (Sp5M) with RanGAP1-SUMO2 (RG-S2) in a 1:1 ratio in the buffer containing 5mM NaCl, 20mM Tris, pH 8.0 and 1mM β-mercaptoethanol. We purify this mixture by the Gel filtration (Superdex 200, GE Healthcare) and also confirmed the interaction between these two proteins because they seem to elute together in one single peak (Gel Filtration data not shown). When we run a SDS-PAGE gel to check the eluted peak, we can observe that they are coming off together and the stoichiometric proportion is almost 1:1, which means that the complex has been formed successfully (Figure 32).

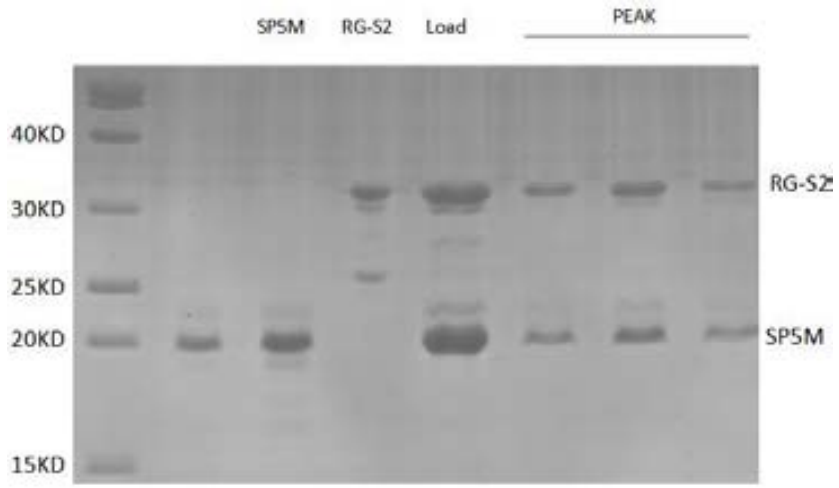


Figure 32 – The Complex-SENP5-CD-C712S and RanGAP1-SUMO2 formation and purification. The inactive SENP5-CD-C712S is mixed with RanGAP1-SUMO2 as the 1:1 ratio. After the gel filtration purification the peak running in the SDS-PAGE gel shows that they elute together.

2.4 Crystallization of SENP5 in complex with the SUMO2 precursor and with the RanGAP1-SUMO2 substrate

Since we are able to produce and purified the complexes of SENP5-CD-C712S with SUMO2 precursor (Sp5-S2p) and with RanGAP1-SUMO2 (Sp5-RG-S2), we next set up crystal trays of these two complexes, and also of the single SENP5-CD, to check whether they could form crystals. All proteins were sprayed into 96-well sitting drop plates, mixed with commercial screen conditions and stored at 18°C. We could only found crystals of Sp5-S2p showing up in a few days in some conditions (Table 8 and Figure 33), whilst we didn‘t observe any crystal of Sp5-RG-S2 and of the single SENP5-CD. Next we optimized the conditions in hanging drop 24-well-plates by mixing 1μL of Sp5-S2p with the corresponding buffer and stored them at 18°C. After several days of incubation, bigger crystals showed up, they were fished up and stored in liquid nitrogen. The X-ray diffraction was carried out at the Alba synchrotron. Unfortunately we could only get a weak diffraction signal around 7Å, which makes impossible to solve the structure (Diffraction data in Table 9). We also tried different methods, including dehydration, to improve the low resolution of the crystals, but they proved to be unsuccessfully.

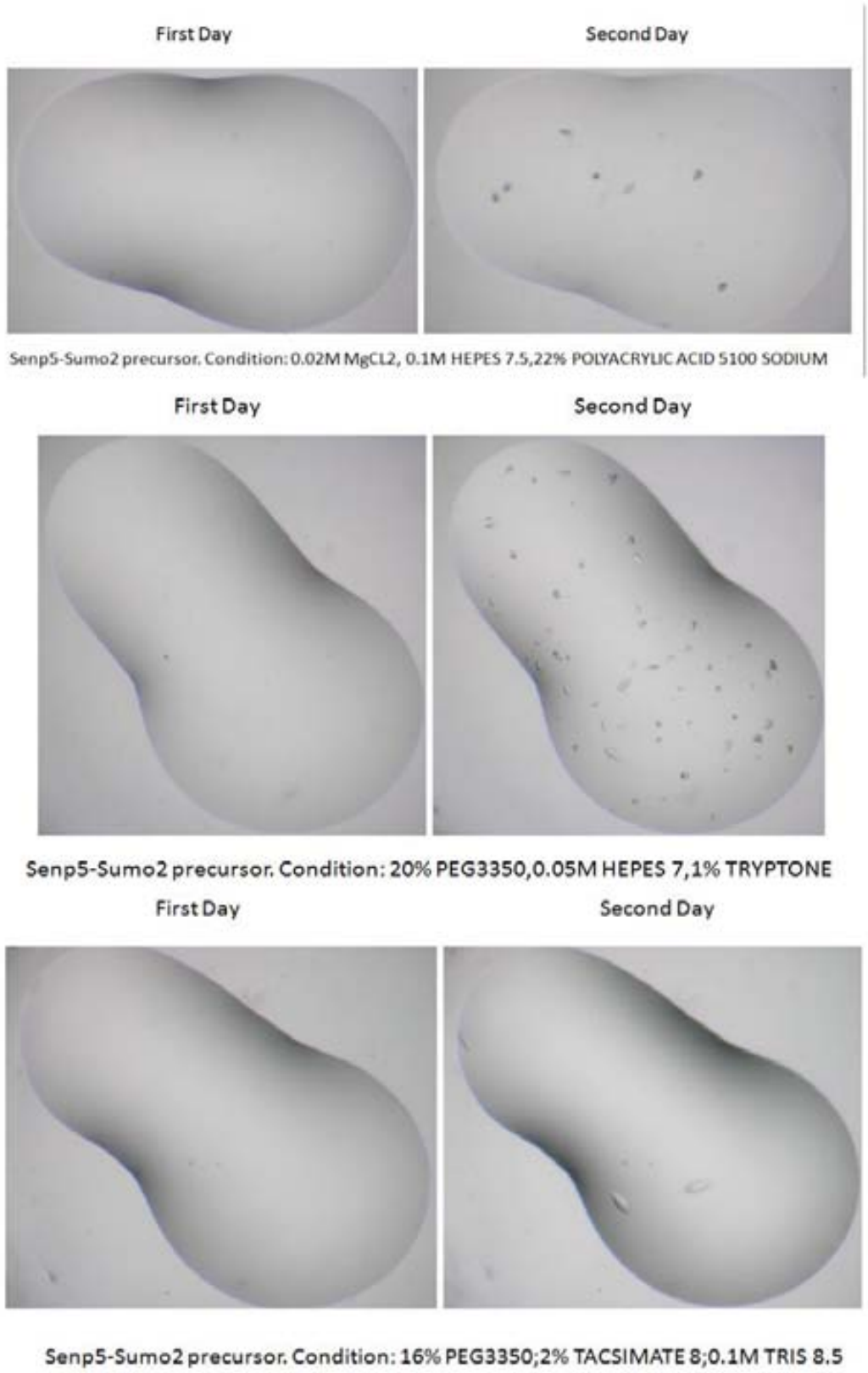
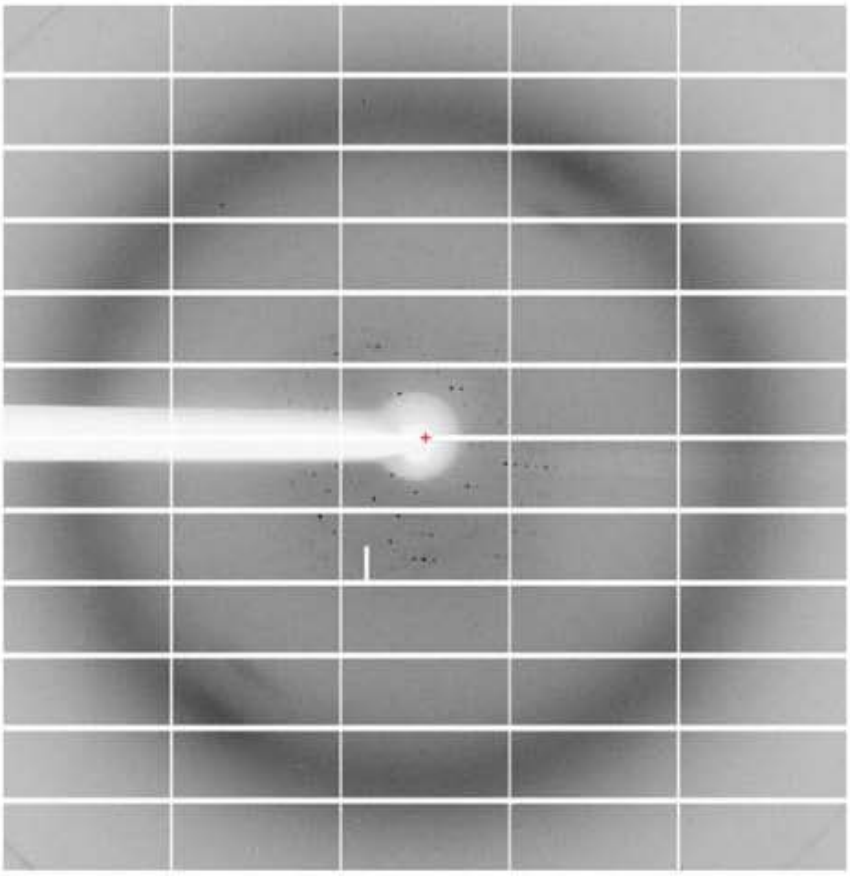


Figure 33 – Crystals of Sp5-S2p.

Screens	Conditions	Showing up
PACT	0.2M $MgCl_2$, 0.1M Tris PH 7.5, 10% PEG8000	2nd Day
PACT	0.2M Potassium Citratete, 20% PEG3350	2nd Day
Proplex	0.1M Tris PH 8.5, 20% PEG6000	2nd Day
Index	0.1M HEPEs PH 7, 5% Tacsimate PH 7, 10% PEG5000MME	2nd Day
Index	0.1M Bis-Tris PH 5.5, 25% PEG3350	2nd Day
Index	0.2M $MgCl_2$, 0.1M Bis-Tris PH 6.5, 25% PEG3350	2nd Day
Index	0.2M Tri-Sodium Citrate, 20% PEG3350	2nd Day
Index	0.1M Potassium thiocyante, 30% PEG2000MME	2nd Day
Index	0.2M NH_4Acetate, 0.1M Bis-Tris PH 5.5, 25% P3350	2nd Day
PEG/ION	0.1M Tris PH 8.5, 5% Tacsimate PH 8, 16% PEG3350	2nd Day
PEG/ION	0.05M HEPEs PH 7, 1% Trypton, 20% PEG3350	2nd Day
JCSG PLUS	0.2M $NaCl_2$, 0.1M HEPEs PH 7.5, 10% Iso-Propanol	2nd Day
SS1&2	0.4M NH_4H_2PO4	2nd Day

SS1&2	Imidazole PH 7	2nd Day
Morpheus	0.09M NPS, 0.1M Sys1 PH 6.5, 30% EDO_P8K	2nd Day
Morpheus	0.09M NPS, 0.1M Sys1 PH 6.5, 30% GOL_P4K	2nd Day
Morpheus	0.09M NPS, 0.1M Sys2 PH 7.5, 30% P550MME_P20K	2nd Day
Morpheus	0.06M Divalents, 0.1M Sys1 PH 6.5, 30% EDO_P8K	2nd Day
Morpheus	0.06M Divalents, 0.1M Sys2 PH 7.5, 30% GOL_P4K	2nd Day
Jena Classic II	0.1M HEPEs PH 7.5, 5% Iso-Propanol	2nd Day
Jena Classic II	0.05M NaAcetate PH4.6, 60% Ethanol, 15% PEG6000	2nd Day
Jena Classic I	0.2M CaAcetate PH4.6, 0.1M HEPEs PH 7.5, 10% PEG8000	2nd Day
Jena Classic I	0.5M KCL, 10% Glycerol, 12% PEG8000	2nd Day
Jena Classic I	0.1M Mes PH6.5, 10% PEG20000	2nd Day
Jena Basic	0.1M $LiSO_4$, 15% PEG8000	2nd Day

Table 8- The screens and conditions of Sp5-S2p crystals.



Data Collection	Sp5-Sp2
Space group	P3
a, b, c (Å)	133.67, 133.67, 96.23
α, β, γ (°)	90, 90, 120
Resolution (Å)[a]	50 – 7.5 (7.88– 7.15)
R_{merge}[b]	0.148
I/σ_I	8.71
Completeness (%)	95.7 (95.2)

Table 9- Sp5-S2p crystal diffraction pattern and data set characterization.

Discussion

1. Discussion of the USP part.

Our *in vitro* assays indicate that USP28 does not display strict chain-type specificity for either K48 or K63-linked poly-ubiquitin chains, showing only subtle preferences in their de-ubiquitinating activity. In contrast to the other DUBs families like the OTU DUBs (Mevisen et al., 2013), most of the members of the USP family have been reported to display a promiscuous de-ubiquitinating activity preference in an assay against the eight possible di-ubiquitin linkages (Faesen et al., 2011). An exception to this is the case of the tumor suppressor CYLD, which possesses specific de-ubiquitinating activity for K63-linked ubiquitin chains synthesized in response to cytokine-mediated activation of TRAF2 and TRAF3 ubiquitin E3 ligases and thus antagonizing the NF-kB signaling (Komander et al., 2008). However, in our *in vitro* assays with all possible di-ubiquitin linkages, we only observe de-ubiquitinating activity against K48, K63 and K11-linked di-ubiquitin substrates, displaying a comparable activity for all our truncated USP28 constructs. Interestingly, these are the three reported main types of di-ubiquitin linkages with a clear connection to cellular functions (Komander and Rape, 2012), and in particular K11-linked poly-ubiquitin chain has been described as an alternative degradation signal in the cell cycle progression (Wickliffe et al., 2011). It would be interesting to check the structural determinants for the preference cleavage of USP28 for these particular chains, which are not dependent on outer domains, and as described in CYLD and in some OTU DUBs members, it might depend on specific interactions with the catalytic domain.

Despite this substrate preference in USP28 in the di-ubiquitin array, our *in vitro* assays with the USP28 truncated constructs indicate that the role of the putative ubiquitin-binding domains in the N-terminal region is not immediately clear. The NMR structure indicates that the USP28 N-terminal region is mostly disordered and only the 3-helix bundle of the UBA domain appears to be a compact globular domain. In our *in vitro* assays the removal of the N-terminal region, that includes the UBA, UIM and SIM domains, does not impair the de-ubiquitinating activity of USP28. Thus the interaction of ubiquitin chains with USP28 during catalysis is not strictly controlled by the interaction with the ubiquitin binding domains in the N-terminal region. This could also explain the lack of discrimination between the K48 and K63-linkage poly-ubiquitin chains in our de-ubiquitinating assays.

Interestingly, although loss of the N-terminal region did not affect activity in our *in vitro* de-ubiquitinating assays, our results indicate that the SUMO modification (mostly on Lys99) of the N-terminal region strongly compromises the activity of USP28. SUMO conjugation can result in different outcomes, including the modification of the enzymatic activity of the target substrate that is modified. Another example of regulation of an enzymatic activity by SUMO modification includes DNA mismatch repair protein TDG (Thymine DNA Glycosylase), where SUMO conjugation reduces its affinity to DNA and promotes the TDG release form the a basic site (Baba et al., 2005). In USP28 the SUMO modification on Lys99 could result in a change in the orientation of the N-terminal region that precludes the interaction of the ubiquitin substrates on the active site of the protease. This inhibition

of de-ubiquitinating activity would not depend on the type and length of the ubiquitin chain, since similar results can be observed for both K48 and K63-linked substrates. Structural studies with SUMO conjugated to USP28 would shed some light to this negative regulation of USP28 by SUMO.

Our experiments indicate that the inhibition of the USP28 catalytic activity only occurs when SUMO is covalently linked to the N-terminal region, forming an isopeptidic bond with Lys99. In our *in vitro* assays we have not observed any non-covalent SUMO inhibition of USP28 activity, even at high concentrations of SUMO and in the presence of the putative SUMO-Interaction motif (SIM) that is identical to that described for USP25. Additionally, we have also not detected interactions between SUMO1 or SUMO2 and the N-terminal region of USP28 using size-exclusion chromatography (data not shown), potentially due to weak-binding affinity between both proteins. In USP25 this non-covalent interaction with the SIM was proposed to promote SUMO conjugation to the N-terminal region in the absence of any SUMO E3 ligase, by facilitating interaction and transfer from the E2-SUMO-thioester conjugating enzyme (Meulmeester et al., 2008). This SUMO conjugation mechanism in the absence of E3 ligase by means of the SIM domain has also been postulated in other examples such as the BLM protein (Bloom‘s syndrome RecQ helicase)(Zhu et al., 2008). The presence of an identical SIM region indicates that in USP28, a similar SUMO conjugation mechanism would be expected. However, we do not observe any SUMOylation of USP28 in the absence of an E3 ligase activity and have not seen an effect of high non-covalent SUMO levels on activity in our

assays. As the SIM is present in a region that appears disordered in the NMR structure, it is possible that the difference is due to sequence divergence between USP25 and USP28 in this region or it is regulated by post-translational modifications. In USP28, Serine 67, that is directly C-terminal to the UBA domain, is phosphorylated following DNA damage and this residue is not conserved in USP25. We speculate that this phosphorylation event could affect the structure in a way that could influence SUMO or ubiquitin binding.

In summary, the SUMOylation of the N-terminal region of USP28 impairs its de-ubiquitinating activity, similarly to what has been reported for USP25. However, we find that in contrast to USP25, the UIM domains of USP28 are not critical for its activity on ubiquitin chains and we do not see any evidence for non-E3 mediated SUMOylation of USP28 *in vitro*. These results add additional detail to the extensive crosstalk between SUMOylation and ubiquitination mechanisms and suggest that active SUMOylation of USP28, and perhaps many other USPs, may influence the half-life of their substrates. USP28 has been implicated in the regulation of c-Myc stability as well as the DNA damage response (DDR) through its interactions with 53BP1 and its identification as a substrate for damage induced phosphorylation by the ATM/ATR kinases. Both Myc signaling and the DDR are extensively regulated by ubiquitin and sumoylation, however the loss of USP28 *in vivo* impairs Myc driven adenocarcinomas but does not have a strong impact on the DDR. Further work will be required to elucidate the precise roles of USP28 and its regulation by SUMOylation.

2. Discussion of the SENP part.

SENP3 and SENP5 constitute a subfamily of SENPs that share considerable sequence homology and exhibit similar substrate specificities. Both of them are localized at the nucleolus and they also have specificity for SUMO-2 and SUMO-3 but less so for SUMO-1. However, the structures of these two proteins are still unknown. One of the reasons is that these two proteins are quite hard to be expressed in *E.coli*. We tried many ways (different vectors, fusion tags, expression strains and conditions) to express different catalytic domains of SENP3/5 but most of them failed. At last only the SENP5-NΔ568 construct could be expressed soluble in *E.coli* and the SENP5 catalytic domain was proven to be active.

Since the SENP family has quite conserved catalytic domain among different members and the SENP1 and SENP2 is expressed in *E.coli* in high yields, we have been very interested in the crystal structure of members of the SENP3/5 sub-family. Unfortunately, we have tried to crystallize many different combinations, the single SENP5-CD and the complexes of SENP5-CD-C712S with both the SUMO2 precursor (Sp5-S2p) and with the RanGAP1-SUMO2 (Sp5-RG-S2). We only could find crystals of Sp5-S2p in a few different conditions, which we were able to reproduce in bigger drops. Despite the crystals were huge and beautiful, the X-ray diffraction was weak, even after checking many different crystals from different conditions with different treatments, the diffraction is only could reach up to 7Å resolution which is hard to resolve the structure at atomic level.

In any case, we have been successfully got crystals of the complex between SENP5-CD-C712S with the SUMO2 precursor (Sp5-S2p) and figured out several conditions that could reproduce these crystals. We have also have characterized the space group and we have been able to collect a full data set for some of them, although in a very low resolution. Perhaps in the future, by finding novel crystallographic conditions, will be the way to finally get a good diffraction and solve the structure of SENP5 in complex with its substrate.

Back up to the SENP5-CD activity, it has been reported that SENP3/5 have specificity for SUMO-2 and SUMO-3 but less so for SUMO-1 (Gong et al. 2006). In our *in vitro* experiments, we have observed similar results, especially in the de-sumoylation reaction, in which SENP5-CD seems prefer SUMO2/3 over SUMO1. These results might be related to the cellular sub-localization of the SENP3/5 and SUMO2/3, which both are reported located at the nucleolus (Gong et al. 2006).

In summary, the sub-family of SENP3/5 is hard to be expressed in *E. coli* and so far the crystals produced, no matter the single protein of SENP5 or in complex with substrate, are not diffracting well. The SENP3/5 subfamily is a tough but interesting topic, which needs more efforts on it in the future. Hopefully, with what have already been achieved in this work -the production of soluble SENP5-CD, the formation of the complex between SENP5 and SUMO2, the conditions for growing crystals and the low resolution diffraction data- we or other scientists will be able to de-code the structural secrets of SENP3/5 sub-family in the future.

Conclusions

1. Conclusions of the USP part

1.1 High-level expression and purification of the N-terminal region of USP25 and USP28 de-ubiquitinases.

1.2 The two SUMO isoforms, SUMO1 and SUMO2/3 do not display a non-covalent interaction with the N-terminal region of USP28 and USP25 in gel filtration co-elution analysis.

1.3 Identification of the USP28 SUMOylation primary sites of the N-terminal region of USP28 by Mass Spectrometry. Lys99 is the major SUMOylation site, followed by Lys63, Lys85 and Lys115.

1.4 Mutagenesis analysis of the SUMOylation primary sites of the N-terminal region of USP28, by SUMOylation assays using IR1 and Nse2 SUMO E3 ligases, confirm the MS results.

1.5 Crystallization and preliminary X-ray diffraction of the N-terminal domain of the USP28. The crystals diffracted beyond 2 Å resolution but were very anisotropic and the structure could not be solved.

1.6 Recombinant production in *E.coli* and characterization of the proteolytic activity of different constructs of the Catalytic Domain of USP28 and USP25 with polyUbiquitin substrates.

1.7 The N-terminal region removal from USP28 catalytic domain does not seems to affect the *in vitro* proteolytic activity against K48 and K63 poly-Ubiquitin substrates.

1.8 USP28 shows Ubiquitin-linkage specificity for K11, K48 and K63 di-Ubiquitin substrates among the eight possible di-Ub substrates linkages.

1.9 SUMO modification on USP28 decreases the activity of the protease using K48 and K63 poly-Ubiquitin substrates analysis. These results suggest a potential regulation of the USP28 activity by SUMO conjugation on its N-terminal domain.

1.10 Identification of the USP28 SUMOylation primary sites of the Catalytic Domain of USP28 by Mass Spectrometry. Lys99 is again the major SUMOylation site, followed by Lys115, Lys385 and Lys511 and Lys513.

1.11 Non-covalent binding of SUMO2 does not modify the activity of the different USP constructs using K48 and K63 poly-Ubiquitin substrate analysis.

1.12 Crystallization of the catalytic domain of USP25. Unfortunately in all conditions the crystals do not diffract beyond 5Å resolution.

2. Conclusions of the SENP part

2.1 Identification of the boundary domains of the SENP5 catalytic domain competent for *E.coli* soluble expression.

2.2 Expression and purification at high levels in *E.coli* of the catalytic domain of the SUMO protease SENP5.

2.3 Characterization of the proteolytic activity of the catalytic domain of SENP5.

2.4 Processing deSUMOylation assays with SUMO isoform precursors indicate a higher activity for SUMO2 isoform. Role of the C-terminal extension of the SUMO isoforms.

2.5 SUMO deconjugation assays with RanGAP1-SUMO1 and RanGAP1-SUMO2/3 isoforms indicate a higher activity compared to the processing reaction and display also a preference for the SUMO2/3 isoforms.

2.6 Production of the inactive mutant of the catalytic domain of SENP5 C712S.

2.7 Complex formation between SENP5-C712S with SUMO2 precursor and with RanGAP1-SUMO2/3.

2.8 Crystallization of the complex between SENP5-SUMO2p.

2.9 In all tested conditions the crystals do not diffract beyond 5Å resolution.

References

Amerik AY, Hochstrasser M. Mechanism and function of deubiquitinating enzymes. Biochim Biophys Acta 2004; 1695:189–207.

Azuma Y, Tan SH, Cavenagh MM, Ainsztein AM, Saitoh H, Dasso M. Expression and regulation of the mammalian SUMO-1 E1 enzyme. FASEB, 2001, 10, 1825-7.

Baba D, Maita N, Jee JG, Uchimura Y, Saitoh H, Sugasawa K, Hanaoka F, Tochio H, Hiroaki H, Shirakawa M. (2005) Crystal structure of thymine DNA glycosylase conjugated to SUMO-1. Nature 435, 979-982.

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

Bencsath KP, Podgorski MS, Pagala VR, Slaughter CA, Schulman BA. Identification of a multifunctional binding site on Ubc9p required for Smt3p conjugation. J Biol Chem, 2002, 277, 47938-45.

Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD. Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. Cell, 2002, 108, 345-356

Best J.L, Ganiatsas S, Agarwal S, Changou A, Salomoni P, Shirihai O, Meluh PB, Pandolfi PP, Zon LI. (2002) SUMO-1 protease-1 regulates gene transcription through PML. Mol. Cell 10, 843–855

Boddy M.N., Howe K., Etkin L.D., Solomon E., Freemont P.S. 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(1996) 971–982.

Bohren K.M., Nadkarni V., Song J.H., Gabbay K.H., Owerbach D. 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. 29 (2004) 29.

Boname JM, Thomas M, Stagg HR, Xu P, Peng J, Lehner PJ. 2010. Efficient internalization of MHC I requires lysine-11 and lysine-63 mixed linkage polyubiquitin chains. Traffic 11:210–20

Bremm A, Freund SM, Komander D. 2010. Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. Nat. Struct. Mol. Biol. 17:939–47

Bruderer, R., Tatham, M. H., Plechanovova, A., Matic, I., Garg, A. K. and Hay, R. T.

(2011) Purification and identification of endogenous polySUMO conjugates. EMBO Rep. 12, 142–148

Bylebyl G.R, Belichenko I, Johnson E.S. The SUMO isopeptidase Ulp2 prevents accumulation of SUMO chains in yeast, J. Biol. Chem. 278 (2003) 44112– 44120.

Carter S, Bischof O, Dejean A, Vousden KH. 2007. C-terminal modifications regulate MDM2 dissociation and nuclear export of p53. Nat. Cell Biol. 9:428–35
Cadwell K, Coscoy L. Ubiquitination on nonlysine residues by a viral E3 ubiquitin ligase. Science 2005; 309:127–30.

Chau V, Tobias JW, Bachmair A, Marriott D, Ecker DJ, Gonda DK, Varshavsky A. A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. Science. 1989; 243(4898):1576–83.

Ciechanover A, Ben-Saadon R. N-terminal ubiquitination: more protein substrates join in. Trends Cell Biol 2004; 14:103–6.

Ciechanover A, Schwartz AL. The ubiquitin-proteasome pathway: the complexity and myriad functions of proteins death. Proc Natl Acad Sci U.S.A. 1998 Mar 17; 95(6):2727-30.

Clague MJ, Coulson JM, Urbé S. Cellular functions of the DUBs. J Cell Sci. 2012 Jan 15; 125(Pt 2):277-86.

Cooper EM, Cutcliffe C, KristiansenTZ, Pandey A, PickartCM, Cohen RE. 2009. K63-specific deubiquitination by two JAMM/MPN +complexes: BRISC-associated Brcc36 and proteasomal Poh1. EMBO J. 28:621–31

David Reverter & Christopher D Lima. Structural basis for SENP2 protease interactions with SUMO precursors and conjugated substrates. NATURE STRUCTURAL & MOLECULAR BIOLOGY. VOLUME 13 NUMBER 12 DECEMBER 2006

Deng L, Wang C, Spencer E, Yang L, Braun A, You J, Slaughter C, Pickart C, Chen ZJ. Activation of the IkappaB kinase complex by TRAF6 requires a dimeric ubiquitinconjugating enzyme complex and a unique polyubiquitin chain. Cell. 2000; 103(2):351–61.

Denuc A, Marfany G. SUMO and ubiquitin paths converge. Biochem Soc Trans. 2010 Feb;38(Pt 1):34-9.

Di Bacco A., Ouyang J, Lee HY, Catic A, Ploegh H, Gill G. (2006) The SUMO-specific protease SENP5 is required for cell division. Mol. Cell. Biol. 26,

4489–4498

Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains—from structures to functions. Nat. Rev. Mol. Cell. Biol. 10, 659–671.

Doss-Pepe EW, Stenroos ES, Johnson WG, Madura K. 2003. Ataxin-3 interactions with Rad23 and valosin-containing protein and its associations with ubiquitin chains and the proteasome are consistent with a role in ubiquitin-mediated proteolysis. Mol. Cell. Biol. 23:6469–83

Dou, H., Huang, C., Singh, M., Carpenter, P. B., & Yeh, E. T. (2010). Regulation of DNA repair through deSUMOylation and SUMOylation of replication protein A complex. Molecular Cell, 39, 333–345.

Drag M., Mikolajczyk J., Krishnakumar I.M., Huang Z., Salvesen G.S. Activity profiling of human deSUMOylating enzymes (SENPs) with synthetic substrates suggests an unexpected specificity of two newly characterized members of the family, Biochem. J. 409 (2008) 461–469.

Dynek JN, GoncharovT,DueberEC, Fedorova AV, Izrael-Tomasevic A. 2010. C-IAP1 and UbcH5 promote K11-linked polyubiquitination of RIP1 in TNF signalling. EMBO J. 29:4198–209

Eaton E.M., Sealy L. Modification of CCAAT/enhancer-binding protein-beta by the small ubiquitin-like modifier (SUMO) family members, SUMO-2 and SUMO-3, J. Biol. Chem. 278 (2003)33416–33421.

Edelmann MJ, Iphofer A, AkutsuM, Altun M, di Gleria K. 2009. Structural basis and specificity of human otubain 1-mediated deubiquitination. Biochem. J. 418:379–90

Erik Meulmeester and Frauke Melchior. SUMO. CELL BIOLOGY. 2008; Vol 452|10 April.

Ernst R, Mueller B, Ploegh HL, Schlieker C. 2009. The otubain YOD1 is a deubiquitinating enzyme that associates with p97 to facilitate protein dislocation from the ER. Mol. Cell 36:28–38

Faesen AC, Luna-Vargas MP, Geurink PP, Clerici M, Merkx R, van Dijk WJ, Hameed DS, El Oualid F, Ovaa H, Sixma TK. The differential modulation of USP activity by internal regulatory domains, interactors and eight ubiquitin chain types. Chem Biol. 2011 Dec 23; 18(12):1550-61.

Felsenstein, J. (1995). PHYLIP (the PHYLogeny Inference Package). (3.57c edn), University of Washington

Finley D. 2009. Recognition and processing of ubiquitin-protein conjugates by the proteasome. Annu.Rev. Biochem. 78:477–513

Forbes S. A., Bhamra G., Bamford S., Dawson E., Kok C., Clements J., Menzies A., Teague J. W., Futreal P. A., and Stratton M. R. (2008) The catalogue of somatic mutations in cancer (COSMIC). Curr. Protoc. Hum. Genet. 10, 10.11.

Gan-Erdene T. Nagamalleswari K, Yin L, Wu K, Pan ZQ, Wilkinson KD. (2003) Identification and characterization of DEN1, a deneddylase of the ULP family. J. Biol. Chem. 278, 28892–28900

Geiss-Friedlander R, Melchior F. Concepts in sumoylation: a decade on. Nat Rev Mol Cell Biol. 2007 Dec; 8(12):947-56.

Gerlach B, Cordier SM, Schmukle AC, Emmerich CH, Rieser E. 2011. Linear ubiquitination prevents inflammation and regulates immune signalling. Nature 471:591–96

Girdwood D. Bumpass D, Vaughan OA, Thain A, Anderson LA, Snowden AW, Garcia-Wilson E, Perkins ND, Hay RT. (2003) P300 transcriptional repression is mediated by SUMO modification. Mol. Cell 18, 1043–1054

Gong L, Millas S, Maul GG, Yeh ET. (2000) Differential regulation of sentrinized proteins by a novel sentrin-specific protease. J. Biol. Chem. 275, 3355– 3359

Gong L, and Yeh, E.T. (2006) Characterization of a family of nucleolar SUMO-specific proteases with preference for SUMO-2 or SUMO-3. J. Biol. Chem. 281, 15869–15877

Goto E, Yamanaka Y, IshikawaA,Aoki-Kawasumi M, Mito-Yoshida M. 2010. Contribution of lysine11-linked ubiquitination to MIR2-mediated major histocompatibility complex class I internalization. J. Biol. Chem. 285:35311–19

Haglund K, Sigismund S, Polo S, Szymkiewicz I, Di Fiore PP, Dikic I. 2003. Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. Nat. Cell Biol. 5:461–66

Haindl, M., Harasim, T., Eick, D., and Muller, S. The nucleolar SUMO-specific protease SENP3 reverses SUMO modification of nucleophosmin and is required for rRNA processing. (2008) EMBO Rep. 9, 273–279

Hang, J. and Dasso, M. (2002) Association of the human SUMO-1 protease SENP2 with the nuclear pore. J. Biol. Chem. 277, 19961–19966

Han Y, Huang C, Sun X, Xiang B, Wang M, Yeh ET, Chen Y, Li H, Shi G, Cang H, Sun Y, Wang J, Wang W, Gao F, Yi J. SENP3-mediated de-conjugation of SUMO2/3 from promyelocytic leukemia is correlated with accelerated cell proliferation under mild oxidative stress. J Biol Chem,, 2010, 17, 12906-15.

Hay RT. (2007). SUMO-specific proteases: a twist in the tail. Trends Cell Biol. 2007 Aug; 17(8):370-6.

Hecker, C, Rabiller, M, Haglung, K, Bayer, P, Dikic, I. Specification of SUMO1- and SUMO2-interacting motifs. J Biol Chem, 2006, 435, 979-982.

Helle D. Ulrich. Ubiquitin, Mutual interactions between the SUMO and ubiquitin systems: a plea of no contest. Trends Cell Biol. 2005 Oct; 15(10):525-32.

Helle D. Ulrich. Ubiquitin. SUMO and the Maintenance of Genome Stability. DNA Repair. Volume 8, Issue 4, 5 April 2009, Pages 429.

Herbst, A., Hemann, M. T., Tworkowski, K. A., Salghetti, S. E., Lowe, S. W., & Tansey, W. P. (2005). A conserved element in Myc that negatively regulates its proapoptotic activity. EMBO Reports, 2, 177–183.

Hetfeld BK, Helfrich A, Kapelari B, Scheel H, Hofmann K. 2005. The zinc finger of the CSNassociated deubiquitinating enzyme USP15 is essential to rescue the E3 ligase Rbx1. Curr. Biol. 15:1217–21

Hochstrasser M. Origin and function of ubiquitin-like proteins. Nature. 2009 Mar 26; 458(7237):422-9.

Hochstrasser, M. SP-RING for SUMO: new functions bloom for a ubiquitin-like protein. Cell, 2001, 1, 5-8.

Huang C, Han Y, Wang Y, Sun X, Yan S, Yeh ET, Chen Y, Cang H, Li H, Shi G, Cheng J, Tang X, Yi J.. SENP3 is responsible for HIF-1 transactivation under mild oxidative stress via p300 de-SUMOylation. EMBO J, 2009, 18, 2748-62.

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

Huang DT, Walden H, Duda D, Schulman BA. Ubiquitin-like protein activation. Oncogene 2004; 23:1958–71.

Hu M, Li P, Li M, Li W, Yao T. 2002. Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. Cell 111:1041–54

Hunter, T. (2007) The age of crosstalk: phosphorylation, ubiquitination, and beyond.

Mol. Cell 28, 730–738

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

Jackson, PK. A new RING for SUMO: wrestling transcriptional responses into nuclear bodies with PIAS family E3 SUMO ligases. Genes, Dev, 2001, 23, 3053-8.

Jacobs, JJ, van Lohuizen, M. Polycomb repression: from cellular memory to cellular proliferation and cancer. Biochim Biophys Acta, 2002, 2, 151-61.

Jentsch S, Pyrowolakis G. Ubiquitin and its kin: how close are the family ties? Trends Cell Biol. 10 (2000) 335–342.

Jiang, M. R., Li, Y. C., Yang, Y., & Wu, J. R. (2003). C-Myc degradation induced by DNA damage results in apoptosis of CHO cells. Oncogene, 2, 2325–2359.

Johnson E.S., Schwienhorst I., Dohmen R.J., Blobel G. The ubiquitin-like protein Smt3p is activated for conjugation to other proteins by an Aos1p/Uba2p heterodimer, EMBO J. 16 (1997) 5509– 5519.

Johnston SC, Riddle SM, Cohen RE, Hill CP. 1999. Structural basis for the specificity of ubiquitin C-terminal hydrolases. EMBO J. 18:3877–87

Jung Hwa Kim, Sung Hee Baek. Emerging roles of desumoylating enzymes. Biochimica et Biophysica Acta 1792 (2009) 155–162.

Kadoya T, Yamamoto H, Suzuki T, Yukita A, Fukui A, Michiue T, Asahara T, Tanaka K, Asashima M, Kikuchi A. (2002) Desumoylation activity of Axam, a novelAxinbinding protein, is involved in downregulation of b-catenin. Mol. Cell. Biol. 22, 3803–3819

Kagey, MH, Melhuish, TA, Wotton, D. The polycomb protein Pc2 is a SUMO E3. Cell, 2003, 1, 127-37.

Kannouche, PL, Wing, J, Lehmann, AR. Interaction of human DNA polymerase eta with monoubiquitinated PCNA: a possible mechanism for the polymerase switch in response to DNA damage. Mol Cell, 2004, 14, 491-500.

Kerscher O, Felberbaum R, Hochstrasser M. Modification of proteins by ubiquitin and ubiquitin-like proteins. Annu Rev Cell Dev Biol. 2006; 22:159-80.

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. and Praefcke, G. J. K. (2013) Multivalent interactions of the SUMO-interaction motifs in the RING-finger protein 4 (RNF4) determine the specificity for chains of the

small ubiquitin-related modifier (SUMO). Biochem. J. 457, 207–214

Kim, KI, Baek SH, Jeon YJ, Nishimori S, Suzuki T, Uchida S, Shimbara N, Saitoh H, Tanaka K, Chung CH. (2000) A new SUMO-1-specific protease, SUSP1, that is highly expressed in reproductive organs. J. Biol. Chem. 275, 14102–14106

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

Komander D, Lord CJ, Scheel H, Swift S, Hofmann K. 2008. The structure of the CYLD USP domain explains its specificity for Lys63-linked polyubiquitin and reveals a B box module. Mol. Cell 29:451–64

Komander D, Rape M. The ubiquitin code. Annu Rev Biochem. 2012; 81:203-29.

Komander D, Reyes-Turcu F, Licchesi JD, Odenwaelder P,Wilkinson KD, BarfordD. 2009. Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. EMBO Rep.10:466–73

Kuo ML, en Besten W, Thomas MC, Sherr CJ. Arf-induced turnover of the nucleolar nucleophosmin-associated SUMO-2/3 protease Senp3. Cell Cycle , 2008, 21, 3378-87.

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

Lallemand-Breitenbach, V., Jeanne, M., Benhenda, S., Nasr, R., Lei, M., Peres, L., Zhou, J., Zhu, J., Raught, B. and de The, H. (2008) Arsenic degrades PML or PML-RARα through a SUMO-triggered RNF4/ubiquitin-mediated pathway. Nat. Cell Biol. 10, 547–555

Lee G.W., Melchior F., Matunis M.J., Mahajan R., Tian Q., Anderson P. Modification of Ran GTPase-activating protein by the small ubiquitin-related modifier SUMO-1 requires Ubc9, an E2-type ubiquitin-conjugating enzyme homologue, J. Biol. Chem. 273 (1998)6503– 6507.

Lee I, Schindelin H. Structural insights into E1-catalyzed ubiquitin activation and transfer to conjugating enzymes. Cell 2008; 134:268–78.

Lee M.H. Lee SW, Lee EJ, Choi SJ, Chung SS, Lee JI, Cho JM, Seol JH, Baek SH, Kim KI, Chiba T, Tanaka K, Bang OS, Chung CH. (2006) SUMO-specific protease SUSP4 positively regulates p53 by promoting Mdm2 self-ubiquitination. Nat. Cell Biol. 8, 1424–1431

Lemak, A., Yee, A., Houliston, S., Garcia, M., Dhe-Paganon, S., Montelione, G.T., Arrowsmith, C. 2012. NMR solution structure of the N-terminal domain of human USP28. Northeast Structural Genomics Consortium.

Licchesi JD, Mieszczanek J, Mevissen TET, Rutherford TJ, Akutsu M. 2011. An ankyrin-repeat ubiquitin-binding domain determines TRABID's specificity for atypical ubiquitin chains. Nat. Struct. Mol. Biol. 19:62–71

Li SJ, Hochstrasser M. A new protease required for cell-cycle progression in yeast. Nature. 1999 Mar 18; 398(6724):246-51.

Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. A small ubiquitin-related polypeptide involved in targeting Ran-GAP1 to nuclear pore complex protein RanBP2, Cell 88 (1997)97– 107.

Mahajan R, Gerace L, Melchior F, Molecular characterization of the SUMO-1 modification of RanGAP1 and its role in nuclear envelope association, J. Cell Biol. 140 (1998) 259– 270.

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

Matunis M.J., Wu J., Blobel G. SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex, J. Cell Biol. 140 (1998) 499– 509.

McCullough J, Clague MJ, Urb′e S. 2004. AMSH is an endosome-associated ubiquitin isopeptidase. J. Cell Biol. 166:487–92

Melchior F, Schergaut M, Pichler A. (2003) SUMO: ligases, isopeptidases and nuclear pores. Trends Biochem. Sci. 28, 612–618

Meluh PB, Koshland D. Evidence that the MIF2 gene of *Saccharomyces cerevisiae* encodes a centromere protein with homology to the mammalian centromere protein CENP-C, Mol.Biol. Cell 6 (1995) 793– 807.

Mendoza HM, Shen LN, Botting C, Lewis A, Chen J, Ink B, Hay RT. (2003) NEDP1, a highly conserved cysteine protease that deNEDDylates Cullins. J. Biol. Chem. 278, 25637–25643

Messick TE, Russell NS, Iwata AJ, Sarachan KL, Shiekhattar R. 2008. Structural basis for ubiquitin recognition by the Otu1 ovarian tumor domain protein. J. Biol. Chem. 283:11038–49

Metzger M.B., Pruneda J.N., Klevit R.E., Weissman A.M. RING-type E3 ligases: master manipulators of E2 ubiquitin-conjugating enzymes and ubiquitination. Biochim. Biophys. Acta 2014 (1843) 47–60.

Meulmeester, E, Kunze, M, Hsiao, HH, Urlaub, H, Melchior, F. Mechanism and consequences for paralog-specific sumoylation of ubiquitin-specific protease 25. Mol Cell, 2008, 30, 610-19

Mevissen TE, Hospenthal MK, Geurink PP, Elliott PR, Akutsu M, Arnaudo N, Ekkebus R, Kulathu Y, Wauer T, El Oualid F, Freund SM, Ovaa H, Komander D. OTU deubiquitinases reveal mechanisms of linkage specificity and enable ubiquitin chain restriction analysis. Cell. 2013 Jul 3; 154(1):169-84.

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

Miyauchi, Y, Yogosawa, S, Honda, R, Nishida, T, Yasuda, H. Sumoylation of Mdm2 by protein inhibitor of activated STAT (PIAS) and RanBP2 enzymes. J Biol Chem, 2002, 51, 5013-6.

Mossessova E., Lima C.D. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast, Mol. Cell 5 (2000)865– 876.

Mukhopadhyay D, Ayaydin F, Kolli N, Tan SH, Anan T, Kametaka A, Azuma Y, Wilkinson KD, Dasso M. (2006) SUSP1 antagonizes formation of highly SUMO2/3-conjugated species. J. Cell Biol. 174, 939–949

Nijman SM, Luna-Vargas MP, Velds A. A genomic and functional inventory of deubiquitinating enzymes. Cell 2005; 123:773–786.

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

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

Ouyang, J, Shi, Y, Valin, A, Xuan, Y, Gill, G. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. Mol Cell, 2009, 34, 145-54.

Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D. 2003. A proteomics approach to understanding protein ubiquitination. Nat. Biotechnol. 21:921–26

Perry J.J., Tainer J.A., Boddy M.N. A SIM-ultaneous role for SUMO and ubiquitin. Trends Biochem. Sci. 33 (2008) 201–208.

Pichler A, Gast A, Seeler J.S., Dejean A., Melchior F. The nucleoporin RanBP2 has SUMO1 E3 ligase activity. Cell 108 (2002)109– 120.

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

Pichler A, Knipscheer P, Saitoh, H, Sixma, TK, Melchior, F. The RanBP2 SUMO E3 ligase is neither HECT- nor RING-type. Nature Struct, Mol Biol, 2004, 11, 984-991.

Pickart CM. Mechanisms Underlying Ubiquitination. Annu Rev Biochem 2001; 70:503–533.

Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. and Hay, R. T. (2012) Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. Nature 489, 115–120

Popov, N., Herold, S., Llamazares, M., Schu¨lein, C., & Eilers, M. (2007). Fbw7 and Usp28 regulate myc protein stability in response to DNA damage. Cell Cycle, 6, 2327–2331.

Popov, N., Wanzel, M., Madiredjo, M., Zhang, D., Beijersbergen, R., Bernards, R. (2007). The ubiquitin-specific protease USP28 is required for MYC stability. Nature Cell Biology, 7, 765–774.

Poulsen, S. L., Hansen, R. K., Wagner, S. A., van Cuijk, L., van Belle, G. J., Streicher, W.,Wikstrom, M., Choudhary, C., Houtsmuller, A. B., Marteijn, J. A. (2013) RNF111/Arkadia is a SUMO-targeted ubiquitin ligase that facilitates the DNA damage response. J. Cell Biol. 201, 797–807

Ravid T, Hochstrasser M. Autoregulation of an E2 enzyme by ubiquitin-chain assembly on its catalytic residue. Nat Cell Biol 2007; 9:422–7.

Renatus M, Parrado SG, D'Arcy A. Structural basis of ubiquitin recognition by the deubiquitinating protease USP2. Structure 2006; 14:1293–1302.

Reyes-Turcu, F. E., Ventii, K. H., and Wilkinson, K. D. (2009) Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes. Annu. Rev. Biochem. 78, 363–397.

Rhodes, D. R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B. B., Barrette, T. R., Anstet, M. J., Kincead-Beal, C., Kulkarni, P., Varambally, S., Ghosh, D., and Chinnaiyan, A. M. (2007) Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia 9, 166–180.

Ringose, L. Polycomb comes of age: genome-wide profiling of target sites. Curr Opin Cell Biol, 2007, 3, 290-7.

Rodriguez, MS, Dargemont, C, Hay, RT. SUMO-1 conjugation *in vivo* requires both a consensus modification motif and nuclear targeting. J Biol Chem, 2001, 276, 12654-9.

Rytinki MM, Kaikkonen S, Pehkonen P, Jääskeläinen T, Palvimo JJ. PIAS proteins: pleiotropic interactors associated with SUMO. Cell Mol Life Sci, 2009, 66, 3029-41.

Sampson DA, Wang M, Matunis MJ. The small ubiquitin-like modifier-1 (SUMO-1) consensus sequence mediates Ubc9 binding and is essential for SUMO-1 modification. J Biol Chem, 2001, 276, 21664-69.

Sacco JJ1, Coulson JM, Clague MJ, Urbé S. Emerging roles of deubiquitinases in cancer-associated pathways. IUBMB Life. 2010 Feb; 62(2):140-57.

Saridakis V, Sheng Y, Sarkari F. Structure of the p53 binding domain of HAUSP/USP7 bound to Epstein–Barr nuclear antigen 1 implications for EBV-mediated immortalization. Mol Cell 2005; 18:25–36.

Saitoh H, Hinchey J. Functional heterogeneity of small ubiquitinrelated protein modifiers SUMO-1 versus SUMO-2/3, J. Biol. Chem.275 (2000) 6252– 6258.

Saitoh H., Sparrow D.B., Shiomi T., Pu R.T., Nishimoto T., Mohun T.J., Dasso M. Ubc9p and the conjugation of SUMO-1 to RanGAP1 and RanBP2, Curr. Biol. 8 (1998) 121– 124.

Satijn DP, Gunster MJ, van der Vlag J, Hamer KM, Schul W, Alkema MJ, Saurin AJ, Freemont PS, van Driel R, Otte AP. RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. Mol Cell Biol, 1997, 7, 4105-13.

Sato Y, Yoshikawa A, Yamagata A, Mimura H, Yamashita M. 2008. Structural basis for specific cleavage of Lys 63-linked polyubiquitin chains. Nature 455:358–62

Schlesinger DH, Goldstein G. Molecular conservation of 74 amino acid sequence of ubiquitin between cattle and man. Nature. 1975 May 29; 255(5507):423-4

Schuettengruber B, Chourrout D, Vervoort M, Leblanc B, Cavalli G. Genome regulation by polycomb and trithorax proteins. Cell, 2007, 4, 735-45.

Shen Z., Pardington-Purtymun P.E., Comeaux J.C., Moyzis R.K., Chen D.J. UBL1, a human ubiquitin-like protein associating with human RAD51/RAD52 proteins, Genomics 36 (1996) 271– 279.

Shuai, K. Modulation of STAT signaling by STAT-interacting proteins. Oncogene, 2000, 21, 2638-44.

Sobko, A, Ma, H, Firtel, RA. Regulated SUMOylation and ubiquitination of DdMEK1 is required for proper chemotaxis. Dev Cell, 2002, 6, 745-56.

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

Sowa ME, Bennett EJ, Gygi SP, Harper JW. 2009. Defining the human deubiquitinating enzyme interaction landscape. Cell 138:389–403

Sriramachandran A.M., Dohmen R.J. SUMO-targeted ubiquitin ligases. Biochim. Biophys. Acta 2014 (1843) 75–85.

Sun, H. Y. and Hunter, T. (2012) Poly-small ubiquitin-like modifier (polySUMO)-binding proteins identified through a string search. J. Biol. Chem. 287, 42071–42083

Takahashi Y, Mizoi J, Toh-E A, Kikuchi Y. (2000) Yeast Ulp1, a Smt3-specific protease, associates with nucleoporins. J. Bio chem. (Tokyo) 128, 723–725

Tanaka K, Nishide J, Okazaki K, Kato H, Niwa O, Nakagawa T, Matsuda H, Kawamukai M, Murakami Y. Characterization of a fission yeast SUMO-1 homologue, pmt3p, required for multiple nuclear events, including the control of telomere length and chromosome segregation, Mol. Cell. Biol. 19 (1999) 8660– 8672.

Tatham, M. H., Geoffroy, M. C., Shen, L., Plechanovova, A., Hattersley, N., Jaffray, E. G., Palvimo, J. J. and Hay, R. T. (2008) RNF4 is a poly-SUMO-specific E3 ubiquitin ligase required for arsenic-induced PML degradation. Nat. Cell Biol. 10, 538–546

Tatham M.H., Jaffray E., Vaughan O.A., Desterro J.M., Botting C.H., Naismith J.H., Hay R.T. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9, J. Biol. Chem. 276 (2001) 35368–35374.

Thrower JS, Hoffman L, Rechsteiner M, Pickart CM. Recognition of the polyubiquitin proteolytic signal. EMBO J. 2000; 19(1):94–102. 7.

Tong, H, Hateboer, G, Perrakis, A, Bernards, R, Sixma, TK. Crystal structure of murine/human Ubc9 provides insight into the variability of the ubiquitin-conjugating system. J Biol Chem, 1997, 34, 21381-7.

Knobel PA, Belotserkovskaya R, Galanty Y, Schmidt CK, Jackson SP, Stracker TH. USP28 is recruited to sites of DNA damage by the tandem BRCT domains of 53BP1 but plays a minor role in double-strand break metabolism. Mol Cell Biol. 2014 Jun; 34(11):2062-74.

Uchiyama K, Jokitalo E, Kano F, Murata M, Zhang X. 2002. VCIP135, a novel essential factor for p97/p47-mediated membrane fusion, is required for Golgi and ER assembly *in vivo*. J. Cell Biol.159:855–66

VanDemark AP, Hofmann RM, Tsui C, Pickart CM, Wolberger C. Molecular insights into polyubiquitin chain assembly: crystal structure of the Mms2/Ubc13 heterodimer. Cell. 2001; 105(6):711–20.

Virdee S, Ye Y, Nguyen DP, Komander D, Chin JW. 2010. Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. Nat. Chem. Biol. 6:750–57

Wang T, Yin L, Cooper EM, Lai MY, Dickey S. 2009. Evidence for bidentate substrate binding as the basis for the K48 linkage specificity of otubain 1. J. Mol. Biol. 386:1011–23

Wang X, Herr RA, Chua WJ, Lybarger L, Wiertz EJ, Hansen TH. Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3. J Cell Biol 2007; 177:613–24.

Weeks SD, Grasty KC, Hernandez-Cuebas L, Loll PJ. 2011. Crystal structure of a Josephin-ubiquitin complex: evolutionary restraints on ataxin-3 deubiquitinating activity. J. Biol. Chem. 286:4555–65

Wickliffe KE, Williamson A, Meyer HJ, Kelly A, Rape M. K11-linked ubiquitin chains as novel regulators of cell division. Trends Cell Biol. 2011 Nov; 21(11):656-63.

Wong BR, Parlati F, Qu K. Drug discovery in the ubiquitin regulatory pathway. Drug Discov Today 2003; 8:746–754.

W. Sheng, X. Liao, Solution structure of a yeast ubiquitin-like protein Smt3: the role of structurally less defined sequences in protein–protein recognitions, Protein Sci. 11 (2002) 1482– 1491.

Wu K, Yamoah K, Dolios G, Gan-Erdene T, Tan P, Chen A, Lee CG, Wei

N, Wilkinson KD, Wang R, Pan ZQ. (2003) DEN1 is a dual function protease capable of processing the C-terminus of Nedd8 deconjugating hyper-neddylated CUL1. J. Biol. Chem. 278, 28882–28891

Xavier Jacq, Mark Kemp, Niall M. B. Martin, Stephen P. Jackson. Deubiquitylating Enzymes and DNA Damage Response Pathways. Cell Biochem Biophys (2013) 67:25–43

Xia ZP, Sun L,Chen X, PinedaG, Jiang X. 2009. Direct activation of protein kinases by unanchored polyubiquitin chains. Nature 461:114–19

Xu P, Duong DM, Seyfried NT, Cheng D, Xie Y. 2009. Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. Cell 137:133–45

Yang, X.J. (2005) Multisite protein modification and intramolecular signaling. Oncogene 24, 1653–1662

Y. Azuma, A. Arnaoutov, M. Dasso, SUMO-2/3 regulates topoisomerase II in mitosis, J. Cell Biol. 163 (2003) 477–487.

Yeh, E.T., Gong L, Kamitani T. (2000) Ubiquitin-like proteins: new wines in new bottles. Gene 248, 1–14

Ye Y, Akutsu M, Reyes-Turcu F, Enchev RI, Wilkinson KD, Komander D. 2011. Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. EMBO Rep. 12:350–57

Zhang, D., Zaugg, K., Mak, T. W., & Elledge, S. J. (2006). A role for the deubiquitinating enzyme USP28 in control of the DNA-damage response. Cell, 126, 529–542.

Zhang H, Saitoh H, Matunis MJ. (2002) Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. Mol. Cell. Biol. 22, 6498–6508

Zu, J., Zhu, S., Guzzo, C.M., Ellis, N.A., Sung, K.S., Choi, C.Y. and Matunis, M. (2008). SUMO binding determines substrate recognition and paralog-selective SUMO modification. J. Biol. Chem. 283, 29405-29415.

Abbreviations

Abbreviations	**Full form**
DTT	Dithiothreitol
SDS	Sodium dodecyl sulfate
PAGE	Polyacrylamide gel electrophoresis
TEMED	N, N, N‘, N‘- Tetramethylethylenediamine
TAE	Tris base, acetic acid and EDTA buffer
TE	Tris and EDTA buffer
Tris	Tris(hydroxymethyl)aminomethane
BSA	Bovine serum albumin
DMSO	Dimethyl sulfoxide
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
GST	Glutathione S-transferase
ATP	Adenosine triphosphate
BME	β-mercaptoethanol
Hepes	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IGEPAL	octylphenoxypolyethoxyethanol
KDa	kilodalton
LB	Luria-Bertani
NMR	Nuclear magnetic resonance
PCR	Polymerase chain reaction
Tween-20	Polysorbate 20
uM	micromolar
mM	milimolar

Summary

Summary in English

USP28 is a member of a family of deubiquitinating enzymes (DUBs), homologous to USP25, belonging to the USP family. USP28 has been involved in IR-induced apoptosis and in the stability of numerous DDR regulators. On the other hand, SENP5 is a member of the Sentrin/SUMO-specific proteases (SENP) in humans, which is reported to be involved in mitosis and/or cytokinesis. In this thesis, the main goal has been to elucidate their crystal structures and to characterize the regulation of their proteolytic activities.

With respect to the USP28, we have produced different constructs of their Catalytic Domain (CD) and N-terminal región (NT), and have characterized their activities against different poly-Ubiquitin substrates. We have also identified the USP28 SUMOylation primary site by Mass Spectrometry (K99). We have been able to obtain nice crystals of the NT of USP28 and of the CD of USP25. Unfortunately, the diffractions of these crystals were weak and the structure could not be solved. We have also identified the di-Ubiquitin chain specificity of USP28 (K11, K48 and K63 linkages). Our *in vitro* analysis indicates that the USP28 proteolytic activity can be regulated by covalent SUMO modification at the N-terminal region; however, the presence of this N-terminal region is not strictly necessary for the USP28 activity. Considering there are one SIM, one UBA and two UIMs in the N-terminal region, this issue would be interesting to be explored deeply in the future.

Regarding to the SENP part, we have put many efforts on the protein expression of different constructs of the Catalytic Domains (CD) of SENP3 and SENP5. Finally we have been able to produce in high yields a soluble construct of SENP5-CD in *E.coli*. We have characterized the SUMO processing and SUMO deconjugation reactions of SENP5-CD. We also have formed complexes between the inactive mutant of the SENP5-CD-C712S with SUMO2 precursor (Sp5-S2p) and with RanGAP1-SUMO2 substrates. We could only get crystals of Sp5-S2p, however the diffraction was so weak that we could not solve the structure at high resolution.

Summary in Catalan

USP28 és un enzim de-ubiquitinasa (DUBs), homologa a USP25, que pertany a la família USPs. USP28 s'ha involucrat en apoptosis induïda per IR i en l'estabilitat de vari regulador de DDR (Reparació de Dany al DNA). D'altra banda, SENP5 és una proteasa especifica per Sentrin/SUMO de la família de SENP, que s'ha descrit que està involucrada en mitosi i citocinesi. En aquesta tesi, el principal objectiu ha estat la resolució de les seves estructures cristal·lines i la caracterització de la regulació de la seva activitat proteolítica.

Respecte a la part de USP28, hem aconseguit produir diferent constructes del seu Domini Catalític (CD) i de la regió N-terminal (NT), i hem caracteritzat les seves activitats sobre diferent substrats poli-Ubiquitina. També hem identificat el lloc de SUMOilització primari de USP28 per Espectrometria de Masses (K99). També hem aconseguit obtenir cristalls del NT de USP28 i del CD de USP25. Desafortunadament, la difracció d'aquests cristalls és dèbil i la seva estructura no s'ha pogut resoldre. També hem identificat l'especificitat de cadena di-Ubiquitin en USP28 (K11, K48 i K63 cadenes). Els nostres assajos *in vitro* indiquen que l'activitat proteolítica de USP28 es pot regular per modificació covalent per SUMO a la regió N-terminal. De totes formes, la presencia de la regió N-terminal no és estrictament necessària per l'activitat USP28. Si considerem que tenim un domini SIM, un domini UBA i dos dominis UIMs a la regió N-terminal, fóra interessant estudiar aquest aspecte amb més detall en el futur.

Respecte a la part de SENP, hem posat molts esforços en la producció d'una construcció soluble del Domini Catalític (CD) de SENP3 i SENP5. Finalment, hem aconseguit produir en grans quantitats un constructe soluble de SENP5-CD en *E.coli*. Hem caracteritzat *in vitro* les reaccions de processament de SUMO i de deconjugació de SUMO. També ham format complexes entre un mutant inactiu de SENP5-CD-C712S amb el precursor de SUMO2 (Sp5-S2p) i amb el substrat RanGAP1-SUMO2. Només hem aconseguit obtenir cristalls de Sp5-S2p, malauradament la difracció era tan dèbil que no hem pogut resoldre l'estructura a alta resolució.