Study of the ubiquitin-like fold in Sde2

Sumanjit Datta

A dissertation submitted for the partial fulfilment
Of BS-MS dual degree in Science.

Indian Institute of Science Education and Research Mohali
April 2018

Certificate of Examination

This is to certify that the dissertation titled '**Study of the ubiquitin-like fold in Sde2**' submitted by Mr. Sumanjit Datta (Reg. No. MS13111) for the partial fulfilment of BS-MS dual degree programme of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Shashi Bhushan Pandit Dr. Kavita Babu Dr. Shravan Kumar Mishra

(Supervisor)

Dated: April 20, 2018

Declaration

The work presented in this dissertation has been carried out by me under the guidance of Dr. Shravan Kumar Mishra at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Sumanjit Datta

MS13111

April 20, 2018

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Shravan Kumar Mishra

(Thesis supervisor)

Acknowledgement

I would like to express my deep gratitude to Dr. Shravan Kumar Mishra, my supervisor for his patient guidance, enthusiastic encouragement, and valuable suggestions. He was always available for discussions, helping and supporting me at all times. He has given me immense freedom in my work and has always encouraged me to think independently and to design experiments on my own. This project would have never been so fruitful without the help and support of my lab members. I would like to mention my special thanks to senior members of the lab Poonam, Prashant, Kiran, Poulami for their guidance and continuous support. Whenever I needed help with any experiment or got stuck they were always there to help me. They were also available for discussing work all the time and it always helped me to get new insights into my work. Next, I would like to thank Rakesh, Karan, Praver, the younger lab members for their enthusiastic participation for discussion about the work. They always shared their experience with any experiments and gave the best ideas when something went wrong. Arashadeep, Asis, Anupa and Bala has been there to help me whenever needed. I would like to thank all of them for making this research experience a memorable one.

I would like to thank my friends Joydip and Devashish for their continuous support in good and bad times.

Sincere thanks to all DBS faculty members, all colleagues and teaching stuffs for their help.

Finally I would like to thank a very special person, Asit Baran Sen, my school teacher and mentor who even after finishing my school days has helped me and supported me. This dissertation is dedicated to him.

Sumanjit Datta

List of figures

List of tables

Notation

UBL	Ubiquitin like protein
ULM	Ubiquitin like modifier
UDP	Ubiquitin like domain protein
DUB	Deubiquitinating enzymes
PTM	Post translational modification
Ub-fold	Ubiquitin like fold
$Sde2_{UBL}$	Ubiquitin like fold in Sde2
Sde2-C	C-terminal of Sde2
FOA	5-fluoroorotic acid

Content

Abstract

Post translational modification by ubiquitin is one of the most studied form of protein modification in eukaryotes. Ubiquitin and ubiquitin like proteins/modifiers (UBL/ULMs) regulate various biological processes by covalently conjugating to proteins. Majority of UBLs have been attributed as a protein modifier like ubiquitin, there are notable exceptions to it as well. Hub1, a ubiquitin like protein regulates RNA splicing but is not involved in protein modification. Another example of ubiquitin like domain protein (UDP) is Sde2, which regulates intron specific pre-mRNA splicing in *Schizosaccharomyces pombe*.

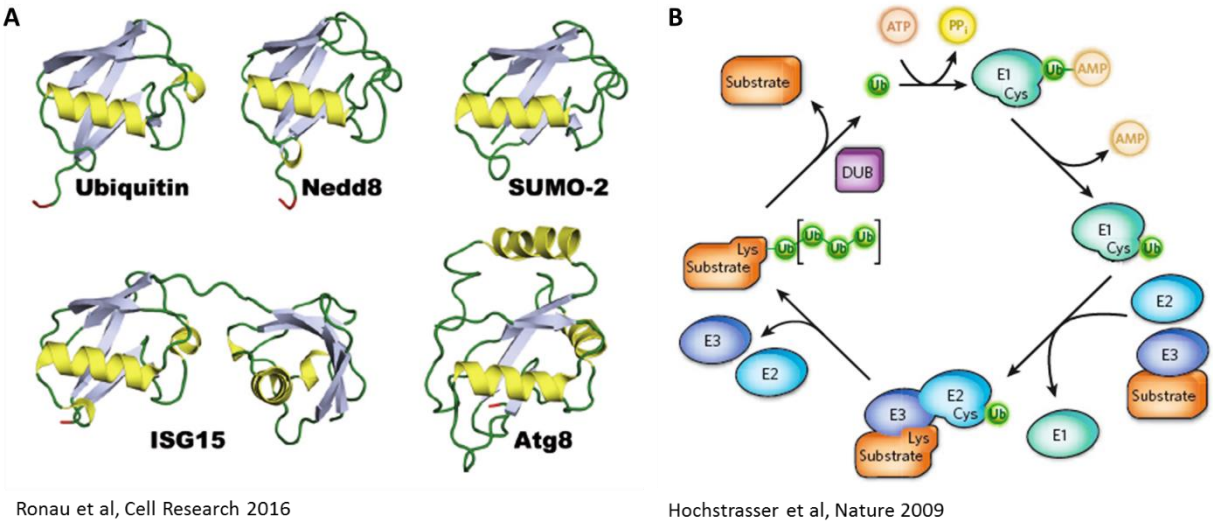
Previously it has been shown that Sde2 has a ubiquitin fold at its N-terminus and a predicted C-terminus rich in helices. Two deubiquitinating enzymes (DUBs) Ubp5 and Ubp15 cleave Sde2 at a conserved GG~KGG motif to generate $Sde2_{UBL}$ and Sde2-C. Despite having low sequence similarity with ubiquitin, Sde2 N-terminus attains a ubiquitin fold and still gets recognized by Ubp5 and Ubp15. This lead to the first part of the study, where we tried to find out the specificity of this DUBs-Sde2 interaction, it seems that the whole UBL fold is necessary for its recognition by Ubp5 and Ubp15 and results in ubiquitin-like processing. Recent reports have elucidated the function of Sde2-C in regulation of RNA splicing in fission yeast and replication stress release factor in mammalian cells; but no significant function has been attributed to $Sde2_{UBL}$. We sought to find out the function of $Sde2_{UBL}$; from our experiment we could see some higher molecular adducts $Sde2_{UBL}$ suggesting a possible role in protein conjugation. Lastly, we hypothesized that presence of UBL fold in Sde2 could regulate its incorporation in the spliceosome but not for other splicing factors like Hub1.

Chapter 1

Introduction

1.1 Ubiquitin and ubiquitin like proteins

Eukaryotic proteins are subjected to various post translational modifications (PTM). Ubiquitination is one of them and in fact it is one of the most prevalent form of modification. Ubiquitin, a small 76 residue long protein, is one of the most conserved protein found in eukaryotes. Structurally it adopts a β-grasp fold which consists of five anti-parallel beta sheets and one alpha helix (Burroughs et al, 2012). Extensive studies have been done on ubiquitin, its molecular mechanism and the diverse nature of ubiquitination is now well characterized. An enzymatic cascade sequentially activates and conjugates ubiquitin to the substrate proteins. Briefly, an E1 enzyme activates ubiquitin in ATP dependent manner and the activated ubiquitin is passed onto the next player of the cascade, E2. Then the ubiquitin is transferred to the substrate by E3 ligases (Kleiger et al, 2014). The diverse nature and the extent of ubiquitination, determines the fate of the substrate protein and regulates its function.



Ronau et al, Cell Research 2016

Hochstrasser et al, Nature 2009

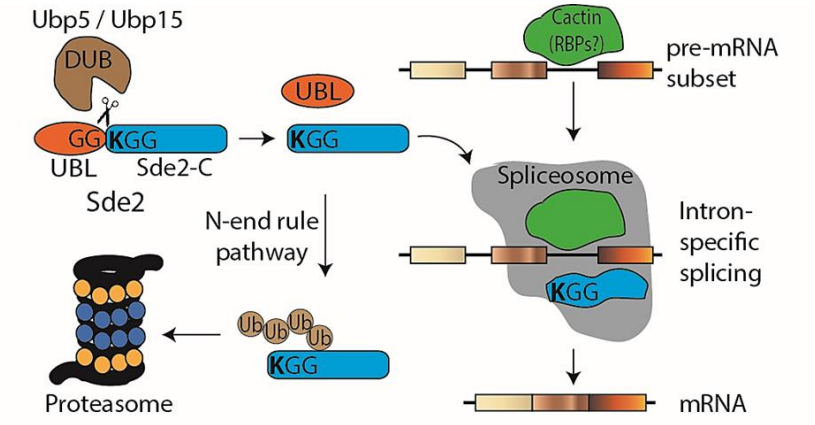
Figure 1.1: **Ubiquitin and UBL and the process of ubiquitination.** A. Structural similarity of ubiquitin and ubiquitin like proteins with β-grasp fold; B. The process of ubiquitination by enzymatic cascade E1, E2, E3 and DUBs.

Over the years a number of proteins have been identified which share a similar three dimensional core structure, known as ubiquitin like proteins (UBLs). They can be divided largely into two groups- type I

UBL or Ubiquitin like modifier (ULM) and type II UBLs or Ubiquitin domain proteins (UDP) (Herrmann et al, 2007; Cappadocia et al, 2018). ULM are those set of proteins which are conjugated to substrate proteins like ubiquitin to a very similar fashion as ubiquitination e.g. SUMO. Nedd8, ISG15. These ULMs also covalently modify proteins by a similar kind of enzymatic cascade involving E1, E2 and E3 enzymes. These modifications regulate a large repertoire of cellular function starting from apoptosis, DNA damage repair, protein degradation, antigen processing to vesicle trafficking, modulation of cell surface receptor and many others. On the other hand, UDPs are single/multi domain protein where ub-fold is present but the protein itself does not get conjugated to other proteins like Rpn10, Hub1 and Esc2. Whilst some are required for ubiquitin recognition such as Rpn10, others like Hub1 has recently been shown to be involved in RNA splicing (Mishra et al, 2011).

1.2 Sde2

Silencing defective 2 (Sde2) was first identified in fission yeast *Schizosaccharomyces pombe* and shown to be involved in genome stability and telomere silencing (Sugiyama et al, 2011). Sde2 harbors an N-terminal ubiquitin like domain, a β-grasp fold (referred as $Sde2_{UBL}$) which is joined to the predicted helical C-terminal domain (referred as Sde2-C) with an invariant motif GGKGG. The putative homologs of Sde2 is found in all eukaryotes from fission yeast to human with a notable exception of *Saccharomyces cerevisiae*, *Candida albicans* and *Pichia pistoris*. Multiple sequence alignment of putative Sde2 homologs shows poorly conserved $Sde2_{UBL}$ at the N-terminus and moderately conserved C-terminal domain. Intriguingly, the linker GGKGG that joins $Sde2_{UBL}$ and Sde2-C is fully conserved across all eukaryotes. We had previously shown that Sde2 gets proteolytically processed by two deubiquitinating enzymes Ubp5 and Ubp15 at GG~KGG motif in *S. pombe* and generate $Sde2_{UBL}$GG and KGGSde2-C. As Sde2 was initially identified as a negative genetic interactor of a splicing factor Hub1, we hypothesized that it might play a role in RNA splicing. We could show that after processing Sde2-C gets incorporated into the spliceosome and mediates intron specific pre-mRNA splicing of some selected introns from a subset of genes (Thakran et al, 2017). Recent study has attributed Sde2 as replication stress release factor in mammalian cells via regulated proteolysis by PCNA (Jo et al, 2016). In another report, Xie et al. have also identified Sde2 in *Arabidopsis* and shown that it plays role in finely coordinating anthocyanin biosynthesis process along with other biological process (Xie et al, 2017).



EMBO J, doi: 10.15252/embj.201796751

Figure 1.2: Schematics of Sde2 function and regulation. Schematic representation of Sde2, its ubiquitin like processing and involvement in pre-mRNA splicing.

1.2.1 Ubiquitin like regulation of Sde2

Compared to other UDPs, Sde2 is unique in many ways- one common factor among them is the presence of ubiquitin fold. Despite having less than 20% sequence similarity with ubiquitin, Sde2 N-terminus attains a ubiquitin fold at the N-terminus. Alignment of amino acid sequence of $Sde2_{UBL}$ with ubiquitin and other ubiquitin like proteins such as Ned8 and SUMO has also shown low sequence homology among UBLs. Yet similar to that of ubiquitin, $Sde2_{UBL}$ also adopts a β-grasp fold. The UBL with the first di-glycine motif in GGKGG makes the $Sde2_{UBL}$ look exactly like ubiquitin. Usually a typical ubiquitin fold is followed by a di-glycine motif which is necessary for its recognition by DUBs and further processing. In case of Sde2, Ubp5 and Ubp15 are able to recognize the UBL and cleave it after the first diglycine motif. This process is similar to the mechanism by which ubiquitin monomer is made from poly-ubiquitin fusion or ubiquitin-ribosome fusion protein (Finley et al, 1989). But this fact that out of 20 DUBs in *S. pombe*, only two of them- namely Ubp5 and Ubp15 recognize the ubiquitin fold of $Sde2_{UBL}$ and process Sde2 to generate the Sde2-C is really intriguing and very unique. This raises the question of the specificity of these two proteases- how do they recognize $Sde2_{UBL}$, is it solely based on the three dimensional structure or by sequence specific surface-surface interaction? Another interesting point is that why are only Ubp5 and Ubp15 capable of this processing? Many studies focused on the substrate specificity of DUBs where the molecular determinants of the specific substrate for each kind of DUBs have been well elucidated (Komander et al, 2009; Ronau et al, 2016). In this context the answer seems to be the interaction resulted from specific interaction with some residues on $Sde2_{UBL}$. But needless to

mention, this distinct and unique recognition of Sde2 by Ubp5 and Ubp15 indicates the flexible nature of substrate recognition and requires more detailed investigation.

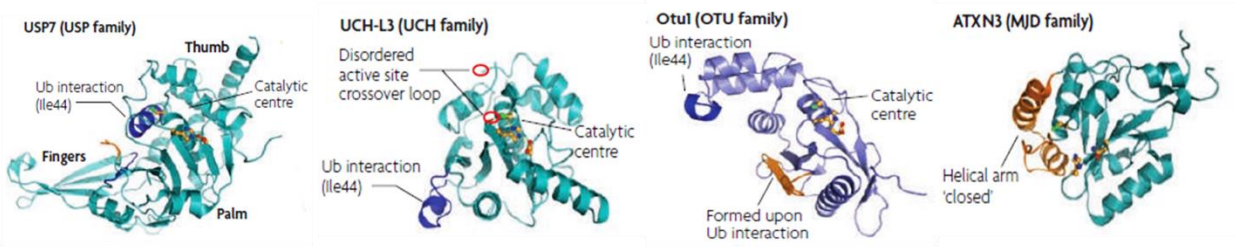
1.2.2 Sde2 processing coupled with RNA splicing

Like its unique mode of processing by two DUBs, Sde2 is novel in its function as well. We have found out that Sde2 is not a general splicing factor but it works as an intron specific splicing factor. Deletion of Sde2 protein from *S. pombe* does not lead to a general splicing defect, instead it shows specific splicing defect. Out of ~5000 introns in *S. pombe*, only a hundred introns of some selected genes gets retained in absence of Sde2 and splicing defect is 1.5 times pronounced compared to wild type. These genes are involved in transcription, cell cycle regulation, chromatin modification, vesicle trafficking etc. covering a wide range of function. As mentioned earlier the linker GGKGG motif is absolutely conserved among all the homologs of Sde2 in eukaryotes; after processing Sde2-C starts with a lysine which is critical for its function. The lysine seems to be required for two purposes- firstly, the N-terminal lysine of Sde2-C is important for proper interaction with other splicing factors such as Cay1/cactin like protein in spliceosome. Secondly, lysine makes the Sde2-C a natural substrate of N-end rule pathway (Bachmair et al, 1986; Varshavsky, 1996), regulating the protein level with relatively shorter half-life. Any defect in processing of Sde2 leads to the accumulation of precursor which is inefficient in spliceosomal incorporation and hence splicing. Changing the N-terminal lysine to any other amino acid affects the splicing and causes growth defect. Thus a ubiquitin like specific cleavage of Sde2 at the conserved GG~KGG motif by Ubp5 and Ubp15 generates ^{Lys}Sde2-C which interacts with splicing factors in the spliceosome and mediates intron-specific splicing.

1.3 Deubiquitinating enzymes (DUBs)

Just like ubiquitination, deubiquitination too plays very important role in cell physiology. Deubiquitinating enzymes (DUBs) are group of proteases that reverse the process of E3 ligases by removing the ubiquitin from substrate proteins by hydrolyzing the isopeptide bond formed between ubiquitin and its targets. Thus they regulate the process of ubiquitination. Apart from that, theses DUBs also cut the peptide bonds between ubiquitin-ribosome precursors and poly-ubiquitin fusion proteins arranged in a head-to-tail manner. They modify the nature of ubiquitination by processing extra ubiquitin from ubiquitinated substrates to ensure the right fate of the proteins. In contrast to mammalian cells having nearly 100 DUBs (Nijman et al, 2005), fission yeast has only 20 DUBs which

belong to one of these four subgroups- USP, UCH, OUT and JAMM (Kouranti et al, 2010). Their diverse nature and regulation has been proposed to specify the linkage pattern of different ubiquitin chain and substrate proteins.



Komander et al, Nature Review, 2009

Figure 1.3: DUBs. Schematic structural representation of different family of deubiquitinating enzymes.

1.3.1 Mechanism of DUBs' specificity

The catalytic domain of DUBs share low sequence homology and harbor distinct structural folds, yet they are all equipped to hydrolyze ubiquitin chains into smaller unit. In general, they all have ubiquitin binding domains where the distal ubiquitin binds and presents its C-terminal Gly to the catalytic center; and the proximal ubiquitin is bound via its ubiquitinated lysine. Upon ubiquitin binding the catalytic site rearrangement takes place and hydrolyze the isopeptide or peptide bonds. Each DUB needs to show the specificity for its substrate ubiquitin or UBL. Some of the specific kinds of regulations have been discussed below-

I. Ubiquitin along with other ULM like Ned8, SUMO, ISG15 are synthesized as a precursor molecule which needs to be cleaved by respective proteases before they could conjugate substrate proteins. There are specific proteases dedicated for each of them, but the mechanisms of by which these proteases recognize UBLs are not very well understood. Even though there are reports suggesting that the last four C-terminal residues preceding Gly-Gly motif is the key for substrate specificity. DUBs recognize Arg-Leu-Arg-Leu-Gly-Gly in case of ubiquitin and hydrolyze the peptide bond in linear fusion of ubiquitin. SUMO, FAT10 share no sequence similarity within these residue and have different proteases for their processing. However, for Ned8, the last stretch Arg-Leu-Ala-Leu-Gly-Gly brings the specificity for deneddylating enzymes; even though report (Gong et al, 2000) indicating dual specificity of USP21 for both ubiquitin and NEDD8 exists. The criticality of these residues might arise while ubiquitin or NED8 upon binding to DUBs extend their C-terminal tail in the catalytic groove, where surface-surface contact stabilizes the interaction followed by

processing. Although a recent study showed that a couple of N-terminal residues of ubiquitin play important role in its recognition by USP2 and changing them to NEDD8 counterpart inhibits processing by USP2 (Shin et al, 2017). Combining these details, it seems that both the terminals along with surface hydrophobic patches (Ile44, Ile36) of ubiquitin are the key regions for its recognition by DUBs and further processing.

II. Unlike phosphorylation, acetylation or other post translational modifications, ubiquitination bears an inherent layer of complexity due to its capacity of getting conjugated to one or more lysine residues on preceding ubiquitin molecule in a poly-ubiquitinated substrate. This determines the topology of the poly-ubiquitin chain as each chain linkages look very different in three dimensional space. Despite this different spatial arrangement, most of the USP family of DUBs do not show any preference for linkage specificity. Evidences supporting both flexibility and specificity by DUBs are documented- USP and OTU family of DUBs have evolved for both lys-48 and lys-63 linkages, but certain JAMM family of DUBs have specificity only for lys-63 linkage (Komander et al, 2009).

III. Ubiquitin can be fused to a substrate both by isopeptide bond with the ε-amino group of acceptor lysine of the substrate proteins or by a peptide bond with the N-terminal amino acid (isoleucine and glycine in case of ubiquitin-ribosomal fusion, methionine in case of poly-ubiquitin fusion in budding yeast). But these two types of ubiquitination are structurally very distinct- isopeptide bond being linked through an elongated flexible hydrophobic chain has rotational freedom whereas the linear peptide chain is rotationally constrained. Due to this reason, not all DUBs are able to cleave peptide bonds between ubiquitin and its corresponding fusion protein. However, data has shown that USP5 and Otulin has activity on poly-ubiquitin fusions whereas USP7 cleaves ubiquitin-ribosome fusion proteins (Grou et al, 2014).

Along with these above mentioned features of determining specificity for DUBs, there are other factors too. Different localization of DUBs limit their activity to specific substrate, likewise post-translational modifications, regulatory adaptor proteins, proteolytic cleavage also contribute to regulate DUB activity and help in determining the specificity. It is apparent from the above discussion that DUBs have preference for their substrate recognition, but there is a lot to unravel about the molecular deterministic factors causing this specificity.

1.4 Aim of the study

From previous observation, we know that two DUBs Ubp5 and Ubp15 cleave Sde2 and generate $Sde2_{UBL}$ and Sde2-C, where Sde2-C gets incorporated into spliceosome and carry out intron specific pre-mRNA splicing. One very intriguing observation was that only Ubp5 and Ubp15 are able to process Sde2 despite of having low sequence similarity with ubiquitin in fission yeast and play direct role in RNA splicing. If DUBs have evolved to process ubiquitin from conjugated substrate or ubiquitin linear fusion, then how do they recognize $Sde2_{UBL}$, what is responsible for this 'dual specificity' behavior of Ubp5 and Ubp15. To be specific we wanted to know the molecular deterministic factor present on the $Sde2_{UBL}$ which makes Sde2 a DUB substrate.

Another interesting aspect was that no obvious function has been assigned to $Sde2_{UBL}$. Previously ubiquitin like protein (Fub1) with a ribosomal protein (S30), arranged in a 'head to tail' manner, have been reported in rat (Olvera et al, 1993). Posttranslational processing of this fusion generates free Fub1 whose function is still unknown, whereas S30 is part of ribosomal small subunit. This is very similar to that of Sde2 where $Sde2_{UBL}$ is fused to a splicing regulator, Sde2-C. With these prelude, we sought to look for the functional significance of $Sde2_{UBL}$- whether it has any regulatory role, or is it just a nonfunctional domain present in Sde2 to generate Sde2-C.

Chapter 2

Experimental methods

2.1 Plasmids and DNA techniques

Schizosaccharomyces pombe strains and plasmids used in this study are listed in Appendix. Preparation of *S. pombe* competent cells, transformation was done following published protocols for *S. cerevisiae* (Knop et al, 1999). All the constructs were made in *S. pombe* vector under either thiamine repressible promoter *nmt81* or constitutive promoter of ENO101 gene. All the constructs have a 3MYC epitope tag at the N-terminus and a 3FLAG epitope tag at the C-terminus. Plasmids with point mutations, insertions or deletions were created by QuikChange site-directed mutagenesis (Agilent) using specific primers. Gene fusions and chimeras were made by overlap extension (SOE) PCR method.

2.2 Growth and complementation assay

For growth and complementation assays, fivefold serial dilution of cells were spotted on indicated agar plates at 30°C and 37°C until growth was observed. The constructs were expressed under the *nmt81* thiamine repressible promoter. The promoter is induced in the absence of thiamine, and 5 µg/ml thiamine was used to repress the promoter. To shuffle-out ura4+ plasmids from *S. pombe*, 5-fluoroorotic acid (FOA) (1.0 g/l of media) was used in agar plates (Thakran et al, 2017).

2.3 Protein expression and western blot

For western blot assays, total protein was isolated from logarithmically growing 1.0 O.D. 600 cells by TCA precipitation and separated on SDS–PAGE and transferred on PVDF membrane. Protein were detected by using anti-MYC (rabbit) and anti-FLAG (mouse) antibody. To induce protein expression from thiamine-repressible *nmt81* promoter, cells were first grown as primary cultures in presence of thiamine. Protein expression was induced by growing the cultures in absence of thiamine for 18–20 h at 30°C (Thakran et al, 2017).

2.4 Random mutagenesis by hydroxylamine

For directed protein evolution, plasmids were isolated in large scale by maxi prep method (Qiagen kit) and 1 µg/µl concentration was obtained. 10 µg of Plasmid was then treated with 1M hydroxylamine in appropriate buffer conditions and incubated at 37°C for 20 hours. Then plasmid was purified and 1 µg plasmid was used per transformation.
(http://www-bcf.usc.edu/~forsburg/hydroxylamine.html)

2.5 Bioinformatics of DNA and protein

Protein sequences were retrieved from NCBI and UniProt database and multiple sequence alignment was done using Clustal omega server. JalView software was used for visualization purpose. Protein structures were retrieved from PDB database, structure was predicted using iTASSER server (Y Zang, 2008). For structural alignment of proteins, Pymol software was used.

Chapter 3

Result

Sde2 is processed by two canonical DUBs Ubp5 and Ubp15 despite sharing low sequence homology at the N-terminus. It is quite intriguing that despite having ubiquitin as their primary substrate, Ubp5 and Ubp15 are able to cleave Sde2. To address this question of dual specificity of factors responsible for Ubp5 and Ubp15's activity on $Sde2_{UBL}$ and ubiquitin, we created a battery of mutations and deletions on UBL and look for the processing of precursor Sde2 in *S. pombe*. Since precursor/processing defective mutant Sde2 is unable to rescue the growth defect of *Δsde2*, we also checked the complementation of those mutated versions of Sde2.

3.1 Sde2 is processed like ubiquitin

Protein 3D structure prediction server iTASSER predicted a ubiquitin like β-grasp fold in Sde2 at the N-terminus with a conserved linker GGKGG motif joining the C-terminal helical domain. Many proteins with features like ub-fold followed by diglycine motif have been categorized as a UBL. Sde2 also follows the same trend. It has a ub-fold flanked by a flexible tail comprising of four amino acids followed by diglycine motif. We reasoned that Ubp5 and Ubp15 cleaves Sde2 at GG~KGG motif starting the Sde2-C with an unusual residue lysine like in case of ubiquitin-ribosome fusion proteins. To show that indeed this is the case, we made alanine and valine substitutions of the terminal glycine which previously have been reported to cause processing defect for ubiquitin-ribosome fusions (Lacombe et al, 2009) and compared with the previously made processing defective mutant of Sde2 (alanine mutations of diglycine motif). Both precursor and processed form of Sde2 was observed in case of Gly84-Ala mutation while the Gly84-Val completely abolishes processing, implying that the processing happens after diglycine motif. With partial processing Gly84-Ala version of Sde2 complements the Sde2 deletion phenotype, but Gly84-Val fails to do so as no processing takes place. With all these observations, we concluded that Sde2 not only gets processed by two DUBs but also follows the similar features of ubiquitin processing.

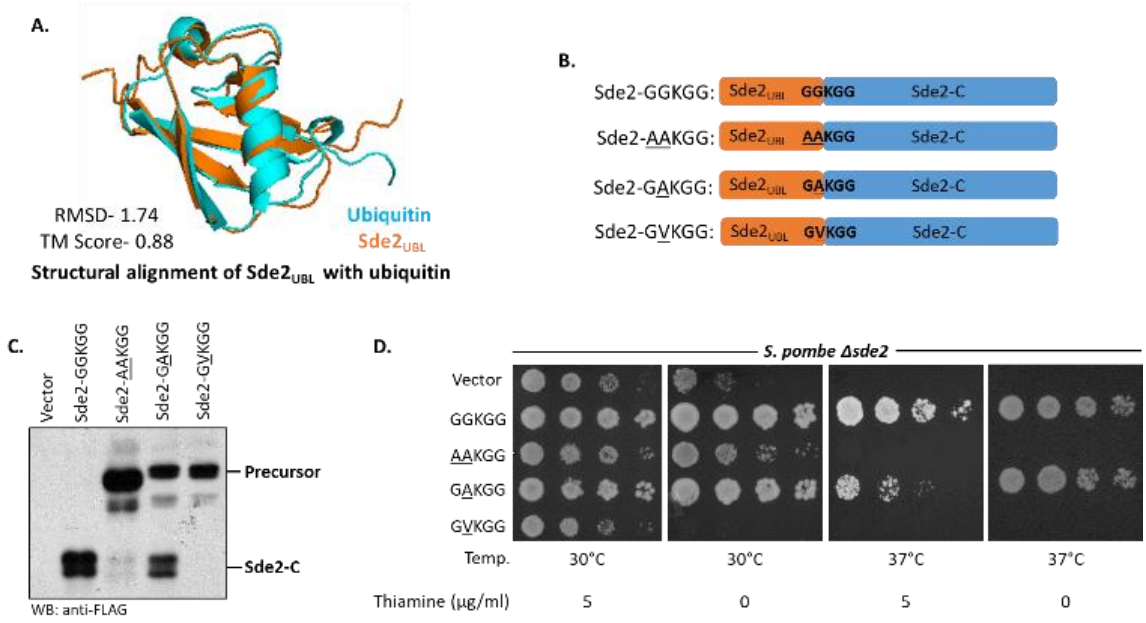


Figure 3.1: Sde2 is processed like ubiquitin. A. Structural alignment of $Sde2_{UBL}$ with ubiquitin; B. Schematic representation of Sde2 and mutants; C. Expression and processing of Sde2 mutants in *S. pombe Δsde2* strain detected by Western blot analysis; D. Complementation of *S. pombe Δsde2* by Sde2 mutants, constructs were expressed under thiamine repressible promoter. Fivefold serial dilution spot assay was done on indicated agar plates and was incubated at 30°C and 37°C until growth appeared.

3.2 Sde2 processing not affected by some UBL fold surface mutations

In order to identify the determinant residues present on $Sde2_{UBL}$, we analysed features of the sequence homology among $Sde2_{UBL}$ and UBLs. As mentioned earlier, $Sde2_{UBL}$ shares less than 20% sequence similarity with ubiquitin or other UBL like protein Ned8 and SUMO. Multiple sequence alignment of amino acids of these proteins showed no significant similarity in the ubiquitin fold region, rather it showed some degree of homology in the last few residues which make the flexible tail of these proteins. Even surprisingly within those four residues preceding the diglycine motif, $Sde2_{UBL}$ has quite a few changes compared to ubiquitin; Sde2 has Thr-Arg-Val-Leu-Gly-Gly residues in its tail instead of Arg-Leu-Arg-Leu-Gly-Gly in ubiquitin. This is the region which upon ubiquitin/UBL binding to DUBs gets extended into the catalytic groove and hydrolysis takes place after the diglycine motif.

We compared sequence of $Sde2_{UBL}$ with SUMO that shares the least similarity among other three and decided to change those residues which are same or similar in $Sde2_{UBL}$, ubiquitin and in Ned8 but different in SUMO. We hypothesized that mutation of those residues on $Sde2_{UBL}$ might show some processing defect of Sde2, if they are important for surface-surface interaction with Ubp5 and Ubp15. We targeted Arg32, Asp37, Arg53, Thr76, Leu77, Arg80 and Val81 spanning the entire length of UBL fold and changed them to the corresponding SUMO counterparts. Site directed mutagenesis (SDM) was performed to change the above mentioned residues in plasmid-borne Sde2 construct (flanked by the epitope tags 3MYC at the N-terminus and 3FLAG at the C-terminus). The processing was observed by western blot analysis; we saw that none of those mutants led to accumulation of Sde2 precursor without visible processed form. Even though precursor gets accumulated in case of Asp37, Leu77 and Arg80 mutants, an ample amount of protein still gets processed which was sufficient to rescue the growth defect in *Δsde2* strain. Therefore, as expected none of these mutant showed any growth defect when transformed in *Δsde2* strain.

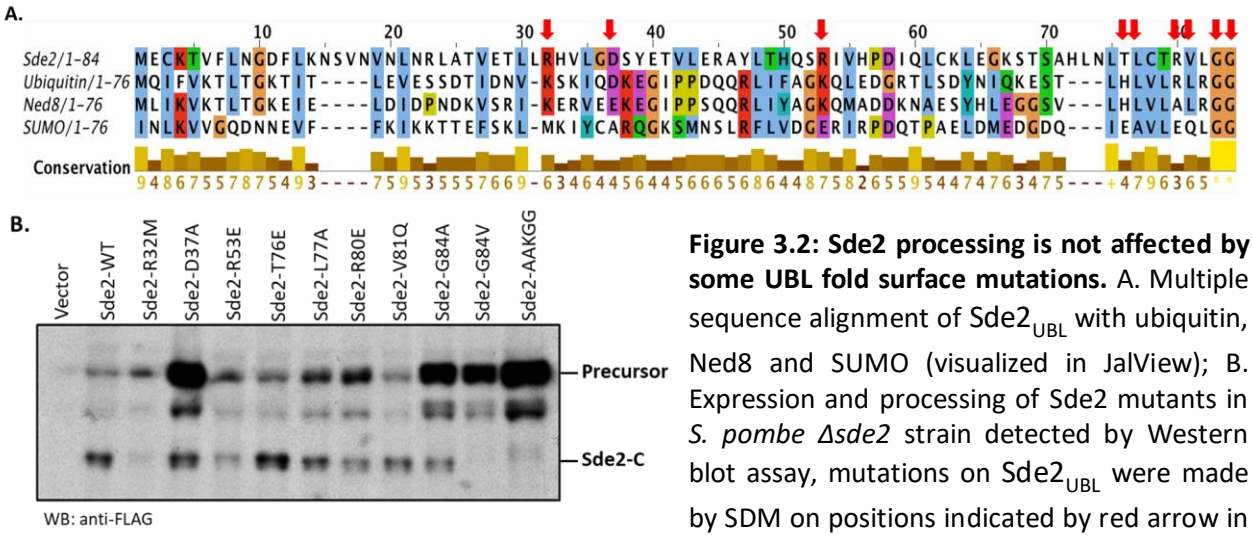


Figure 3.2: Sde2 processing is not affected by some UBL fold surface mutations. A. Multiple sequence alignment of $Sde2_{UBL}$ with ubiquitin, Ned8 and SUMO (visualized in JalView); B. Expression and processing of Sde2 mutants in *S. pombe Δsde2* strain detected by Western blot assay, mutations on $Sde2_{UBL}$ were made by SDM on positions indicated by red arrow in A.

This result suggests that Sde2 recognition by those proteases may not happen solely based on one or two residues, rather a surface or a subdomain might be involved for this purpose.

3.3 Sde2$_{UBL}$ fold is necessary for processing

As single mutations in the Sde2$_{UBL}$ did not show any processing or growth defect, we hypothesized that it could be the whole UBL which is necessary for processing. For this we again carefully analysed the sequence alignment of the Sde2$_{UBL}$ with ubiquitin, NED8 and SUMO and divided the ub-fold into two parts. We deleted the first 72 amino acids of Sde2$_{UBL}$ which shares the least sequence homology with any of the three UBLs starting the protein with 73rd amino acid. Next, we deleted the stretch of 10 amino acids (73rd to 82nd) which corresponds to the flexible and somewhat conserved tail of Sde2$_{UBL}$ with other UBLs leaving the diglycine motif intact. Processing was checked by western blot analysis and it was found that in both truncated versions of Sde2$_{UBL}$, Sde2 processing is affected. This suggests that the ub-fold comprises the necessary surface through which Sde2 interacts with Ubp5 and Ubp15 followed by processing.

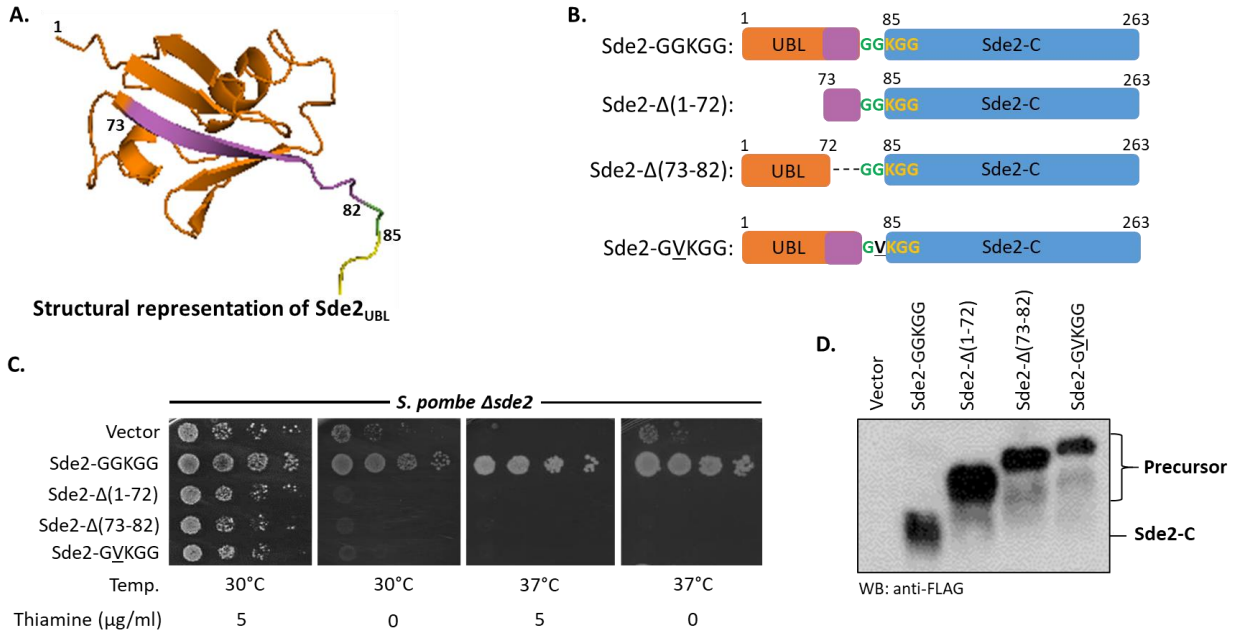


Figure 3.3: Sde2$_{UBL}$ fold is necessary for processing. A. iTASSER predicted Sde2$_{UBL}$ structure, first 72 amino acids are marked in orange, 73rd to 82nd are in magenta, diglycine motif are in green and from 85th onwards in yellow; B. Schematic representation of Sde2$_{UBL}$ with UBL fold deletions and short truncations; C. Complementation of *S. pombe Δsde2* by Sde2 mutants, constructs were expressed under thiamine repressible promoter. Fivefold serial dilution spot assay was done on indicated agar plates and was incubated at 30°C and 37°C until growth appeared; D. Expression and processing of Sde2 mutants in *S. pombe Δsde2* strain detected by Western blot analysis.

Deleting the β-grasp fold makes Sde2 unrecognizable by Ubp5 and Ubp15 which led to processing defect. Whereas for the other short truncation, we speculated that Sde2 with its ub-fold may bind to DUBs but in this case it does not have a flexible tail to present the diglycine motif into the catalytic groove of Ubp5 and Ubp15 and subsequently no processing is observed. As none of these versions are processing efficient, as expected they could not complement the growth defect of Sde2 deletion in *S. pombe*. With this evidence in hand it has now become apparent that the whole UBL fold in Sde2 is necessary for its processing, it is not any single residues which would help in processing. Still importance of some specific residue at very critical position which mediates the recognition has not been ruled out. Similar to that of distal ubiquitin binding to DUBs, $Sde2_{UBL}$ helps in binding Ubp5 and Ubp15 with its β-grasp fold and extends its flexible tail followed by diglycine motif where the catalytic triad in the catalytic groove hydrolyzes the peptide bond between the glycine and the lysine.

3.4 *S. japonicus* Sde2 processing in *S. pombe*

The flexibility shown by Ubp5 and Ubp15 for both Sde2 and ubiquitin as their substrate is quite unique in nature and intriguing as well. To understand the extent of flexibility in substrate recognition exhibited by those two proteases, we looked at the different species of *Schizosaccharomyces* genera and found that four species have been identified- *Schizosaccharomyces pombe*, *Schizosaccharomyces octosporus*, *Schizosaccharomyces cryophilus* and *Schizosaccharomyces japonicus*. Their genome has been fully sequenced and the comparative genomic study for these sister species have been reported (Rhind et al, 2011). We speculated that a similar kind of mechanism where Sde2 ortholog would be getting cleaved by Ubp5 and Ubp15 orthologs exist in those species. Sde2 orthologs in the other species have been identified by BLAST method and multiple amino acid sequence alignment was performed. We found out that the Sde2 orthologs only share 38% sequence homology in their N-term and almost 60% in C-term. This was very surprising, we expected that within the same genus the Sde2 orthologs would share better sequence homology in their N-term, and yet iTASSER predicted structure for the orthologs in all sister species shows presence of a similar ubiquitin like fold. We hypothesized that $Sde2_{UBL}$ in all orthologs are not subjected to evolutionary constrain in terms of sequence but still adopts a ub-fold. The variation in the sequence in those UBL regions indicates a constrain for maintaining the overall three dimensional structure rather than to conserve the sequence similarity.

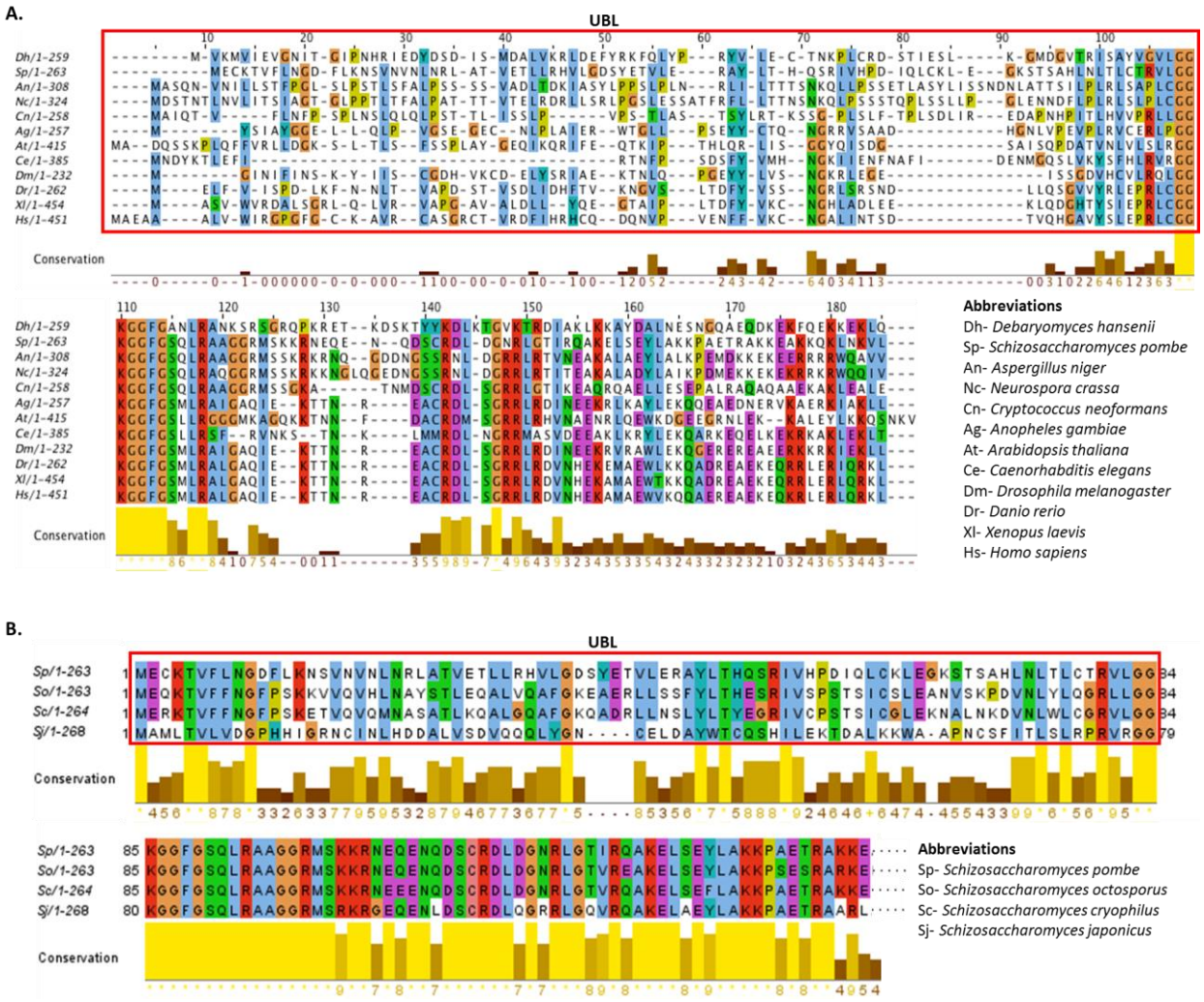


Figure 3.4.1: Sde2$_{UBL}$ is poorly conserved. A. Multiple sequence alignment of Sde2$_{UBL}$ among eukaryotes; B. Multiple sequence alignment of Sde2$_{UBL}$ among the *Schizosaccharomyces* species.

With this in hand, we thought of looking into the processing of Sde2 orthologs from other *Schizosaccharomyces* species in *S. pombe*. Among the other three species, *S. octosporus* and *S. cryophilus* Sde2$_{UBL}$ share almost 60% sequence similarity with *S. pombe* Sde2$_{UBL}$, but for *S. japonicus* it turned out to be only 29%. We speculated that Sde2 from the other two species might get processed due to their reasonable sequence similarity with *S. pombe* and proceeded only with *S. japonicus*. Sde2 from *S. japonicus* (Sj Sde2) was cloned and transformed into *S. pombe* and looked for both processing and complementation. *S. japonicus* Sde2 gets processed in *S. pombe* inspite of low sequence homology, DUBs are able to recognize its ub-fold and cleave at the GG~KGG motif. As expected, *S. japonicus* Sde2 could also complement the growth defect of the Sde2 deletion phenotype, suggesting that the *S. pombe* Sde2 is replaceable with the *S. japonicus* Sde2. Surprisingly, however unlike at 30°C, *S. japonicus* Sde2

could not rescue the growth defect of *Δsde2* in *S. pombe* at higher temperature of 37°C. We initially thought this might arise due to processing defect of *S. japonicus* Sde2 at higher temperature, but when checked we found complete processing and generation of *S. japonicus* Sde2-C. Our best guess is that after processing, Sde2-C of *S. japonicus* could not fully take over the function of *S. pombe* Sde2-C and it is deficient in pre-mRNA splicing at 37°C. It would be interesting to look at the splicing defect of Sde2 dependent targets, whether *S. japonicus* Sde2 is able to recognize and help in splicing of those selected introns.

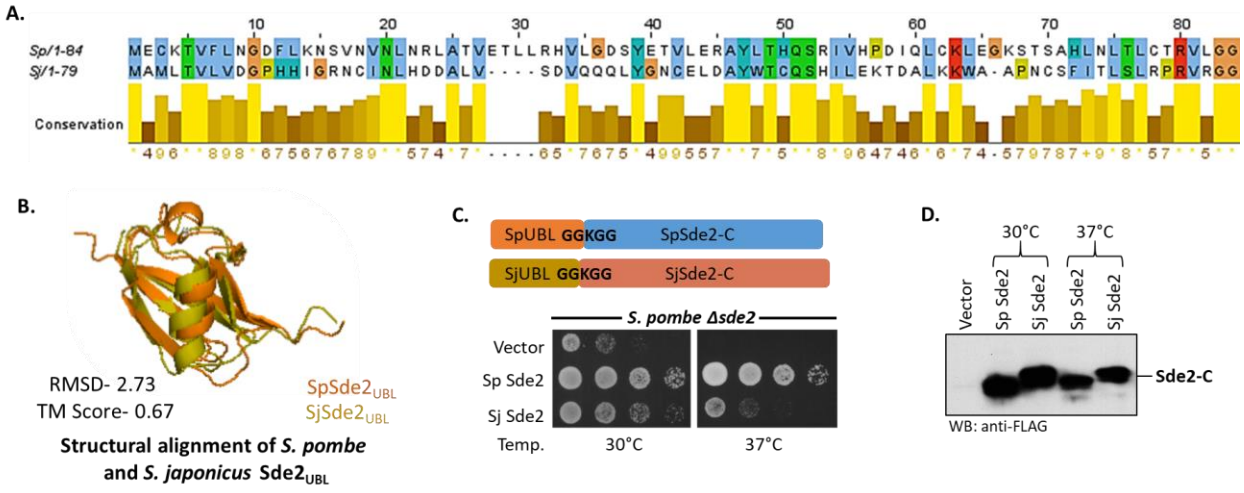


Figure 3.4.2: *S. japonicus* Sde2 gets processed in *S. pombe*. A. Sequence alignment of *S. japonicus* and *S. pombe* $Sde2_{UBL}$ showing only 29% conservation; B. Structural alignment of $Sde2_{UBL}$ of *S. japonicus* and *S. pombe* Sde2; C. Schematic representation of *S. japonicus* and *S. pombe* Sde2 (top), Complementation of *S. pombe Δsde2* by Sde2 of both species (bottom), constructs were expressed under constitutive promoter. Fivefold serial dilution spot assay was done on indicated agar plates and was incubated at 30°C and 37°C until growth appeared; D. Expression and processing of Sde2 mutants in *S. pombe Δsde2* strain detected by Western blot analysis (*S. pombe* Sde2- SpSde2, *S. japonicus* Sde2- SjSde2).

This could be due to improper association with the spliceosome or deficient interaction with other splicing factors. Although it is possible that $Sde2_{UBL}$ of *S. pombe* plays some critical function which becomes indispensable at 37°C and cannot be replaced by *S. japonicus* $Sde2_{UBL}$. Nonetheless, it is worth pointing out that DUBs of *S. pombe* are able to recognize *S. japonicus* $Sde2_{UBL}$ in *S. pombe*, suggesting the evolutionary conservation of ubiquitin like processing of a splicing factor in other fission yeasts.

3.5 Replacement of $Sde2_{UBL}$ with Hub1 does not facilitate processing

$Sde2_{UBL}$ seems to be very unique for getting recognized by two DUBs and show a typical ubiquitin like processing, yet it is replaceable by ubiquitin or NED8 (studied by a PhD fellow Poonam Thakran in our lab). We thought that Ubp5 and Ubp15 then only require a ubiquitin like fold to recognize it as a substrate and cleave Sde2. Ub-$^{Lys}{}_{85}$Sde2-C and NED8-$^{Lys}{}_{85}$Sde2-C fusions were able to complement the growth defect of *Δsde2* strain and gets processed even in *Δubp5 Δubp15* strain (studied by a PhD fellow Prashant Pandit in our lab), suggesting that the other DUBs and deneddylating enzymes are processing the chimeras. This was expected in a way because ubiquitin or NED8 were made as precursor and for them to get functional processing is needed. So, we thought of making a fusion protein with similar sized UDP as $Sde2_{UBL}$ which does not require processing for its function and check whether now Ubp5 and Ubp15 recognize the UBL for processing. One of such UBL which fulfills both the criteria, Hub1 when fused to Sde2-C (Hub1-^{KGG}Sde2-C) did not show any processing and hence no complementation (previously shown by Poonam).

But we reasoned that since Hub1 does not require a processing enzyme, Hub1-^{KGG}Sde2-C chimera is an unnatural substrate for the cell and hence no processing was observed. So, in order to make it more similar like Sde2, we inserted the last 10 residue of $Sde2_{UBL}$ in the chimera (now Hub1-$^{Leu}{}_{73}$Sde2-C) which will provide the UBL from Hub1 and the flexible tail of the inserted amino acid stretch from $Sde2_{UBL}$. When checked for processing, to our surprise we saw no processing of fusion Hub1-$^{Leu}{}_{73}$Sde2-C and no rescue of the *Δsde2* growth defect. This means despite providing both the UBL and the flexible tail followed by the diglycine motif, Ubp5 and Ubp15 are not able to recognize and process chimeric Sde2. This non-processing might arise due to the fact that even though Hub1 shares a ub-fold, it's sequence is very different from $Sde2_{UBL}$ and hence the putative 'critical residues' are missing. In any case we can generalize that the flexibility of Ubp5 and Ubp15 comes only to those UBLs which naturally undergo processing before activation, not to all ub-fold protein. It is also quite possible that the chimera is misfolded or unfolded which lead to accumulation of only precursor form.

In order to gain the activity of Ubp5 and Ubp15 on Hub1-$^{Leu}{}_{73}$Sde2-C chimera by making it more similar to that of WT Sde2, we had a closer look to the structural alignment of $Sde2_{UBL}$ and Hub1. Both the UBLs of Sde2 and Hub1 share the similar β-grasp fold having the secondary structural element arranged in ββαββ manner. From a recent study (Shin et al, 2017), it came to our notice that USP domain containing DUBs also require N-terminal residues of ubiquitin for its recognition along with the flexible tail preceding the diglycine motif.

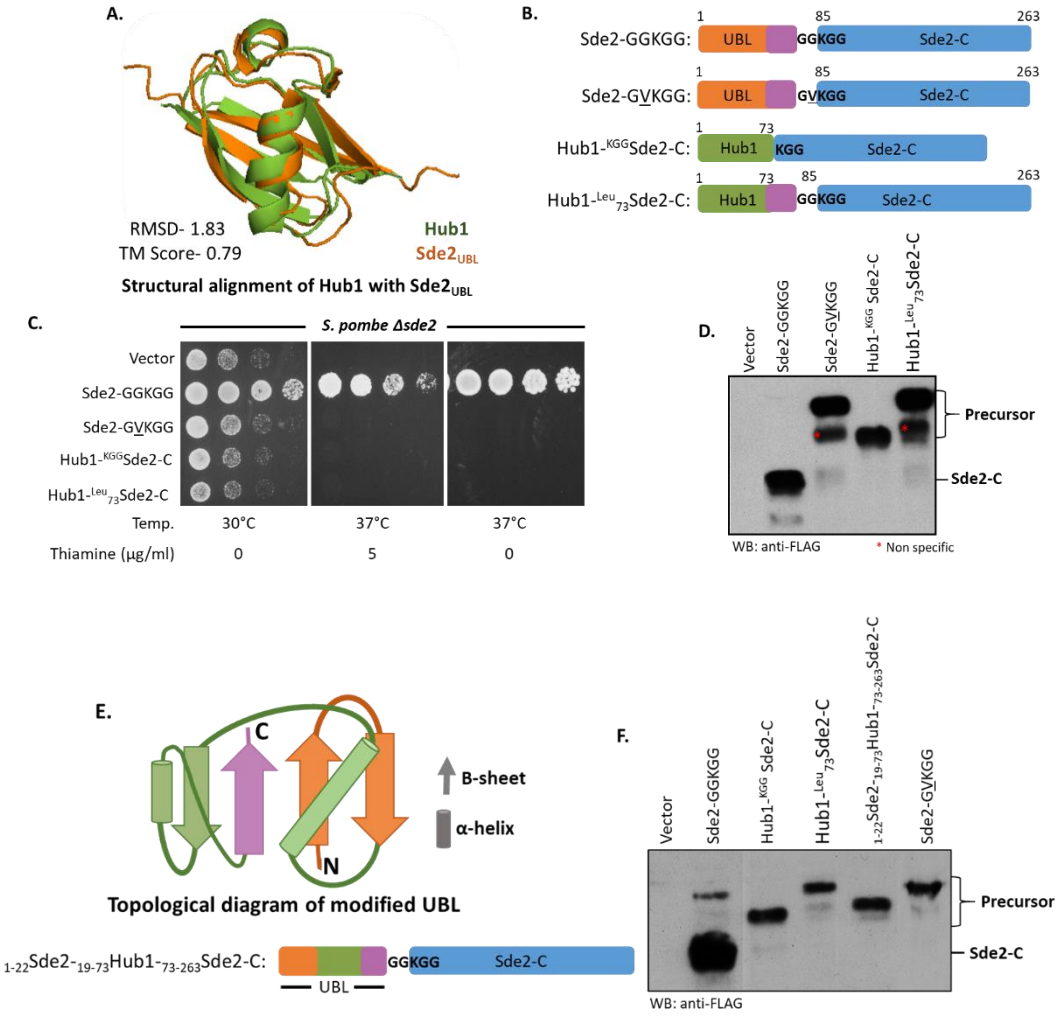


Figure 3.6 Replacement of $Sde2_{UBL}$ fold with Hub1 does not facilitate processing. A. Structural alignment of $Sde2_{UBL}$ with Hub1; B. Schematic representation of Hub1-Sde2 chimeras; C. Complementation of *S. pombe Δsde2* by Hub1-Sde2 chimeras, constructs were expressed under thiamine repressible promoter. Fivefold serial dilution spot assay was done on indicated agar plates and was incubated at 30°C and 37°C until growth appeared; D. Expression and processing of Hub1-Sde2 chimeras in *S. pombe Δsde2* strain detected by Western blot analysis; E. Topological diagram of modified UBL from Sde2-Hub1-Sde2 chimera; F. Expression and processing of new Sde2-Hub1-Sde2 chimeras in *S. pombe Δsde2* strain detected by Western blot analysis.

With this mechanistic insight, we replaced the first two β-sheets of Hub1 in the Hub1-$^{Leu}_{73}$Sde2-C chimera with $Sde2_{UBL}$'s first two β-sheets. We hypothesized that now with this replacement of extreme N-terminal of Sde2, $_{1-22}$Sde2-$_{19-73}$Hub1-$_{73-263}$Sde2-C chimera will have a proper ub-fold where $_{1-22}$Sde2 provides first two β-sheets, $_{19-73}$Hub1 provides next part of β-grasp fold (αββ part) the flexible tail with

diglycine motif and rest of the C-term by $_{73\text{-}263}$Sde2-C. With a hope that now it will get recognized and processed by the DUBs Ubp5 and Ubp15, western blot was performed. But even this chimera was not processed in *S. pombe*.

This suggests that the extreme terminals of Sde2$_{UBL}$ are not sufficient for processing to take place, may be uniqueness in the sequence arrangement of the UBL makes it a novel substrate of Ubp5 and Ubp15.

3.6 Obtaining processing efficient chimera of Hub1-$^{Leu}_{73}$Sde2-C by directed evolution

Up until now we could not figure out the molecular determinants of Sde2 for it to become a DUB substrate. The orthodox rational way of thinking to make a processing defective Sde2 chimera a natural substrate of DUBs did not succeed. So, we took the 'Directed Protein Evolution' approach to circumvent this issue. In this way one can evolve a protein to do somewhat desired function within the laboratory condition. The idea behind this was to create a library of different Sde2 chimeras which will start getting processed, selected by the cells and lead to growth advantage. Since Sde2 deletion at permissive temperature (30°C) is viable yet shows a measurable growth defect, we hypothesized that any processing efficient chimera of Sde2 might get cleaved by the DUBs to generate Sde2-C and result in intron specific pre-mRNA splicing. This would rescue the growth defect and colonies will appear faster than the cells having processing defective chimera. As Sde2 is not an essential gene, in this case the selection would be based on growth rate of cells.

So to start with we took the Hub1-$^{Leu}_{73}$Sde2-C chimera which was processing defective and subjected it to random mutagenesis by addition of hydroxylamine. Hydroxylamine is an inorganic compound which specifically targets the pyrimidine cytosine on DNA. Addition of an extra hydroxyl group to cytosine makes it hydroxlaminocytosine, which can no longer bind to complementary base guanosine, instead it pairs with adenine. So, in next round of DNA replication that DNA stretch will have an A-T base pair replacing the original C-G pair. Thus upon hydroxylamine treatment the whole plasmid will get randomly mutated and incorporate transition mutations at many different places which will have both beneficial and deleterious consequences. Any deleterious changes either in the selection marker, origin of replication sequences or in Sde2-C will lead to growth disadvantage of cell. But if there is any positive mutation in the Hub1 region which now can get recognized by the DUBs and processed, then the cell will have growth advantage.

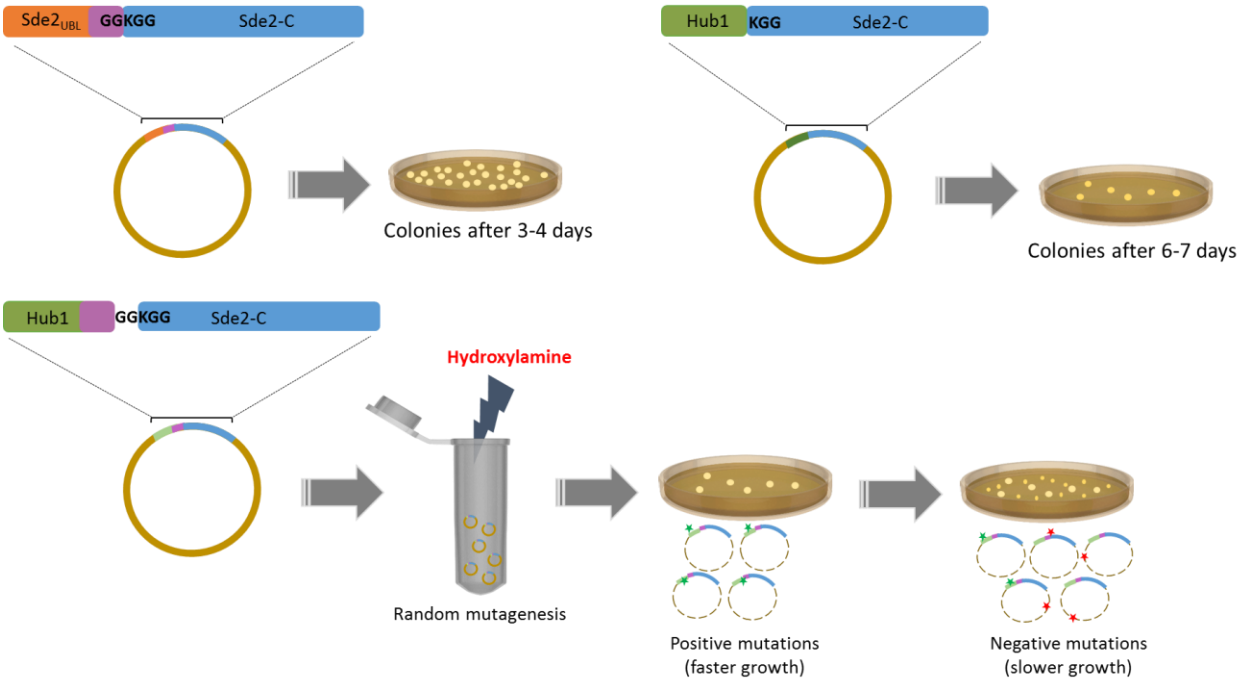


Figure 3.6.1: Directed protein evolution. Schematic representation of Directed protein evolution by hydroxylamine. Using this strategy one can try to evolve a protein for desired optimum function.

Having this strategy in mind, we treated the plasmid having Hub1-$^{Leu}_{73}$Sde2-C chimera with hydroxylamine and transformed into the Sde2 deletion cells and selected those colonies that grew within three-four days. We could compare the growth of those colonies with the WT Sde2 transformed in Sde2 deletion cells. Any colony that came after the initial ones would contain deleterious mutations or missense mutations are not in the focus of the study. After screening around 12 colonies, we finally selected 4 of them which showed the best rescue of growth defect. To confirm that indeed the growth advantage came due to the processing of Sde2 chimera, we performed western blot analysis and saw Sde2-C is getting generated. Next to know the type of mutations that makes the Hub1-$^{Leu}_{73}$Sde2-C chimera processing efficient, the plasmid from those strains was isolated by shuttle prep and sent for sequencing analysis. To confirm the result, we have retransformed those plasmid into Sde2 deletion strain and looked for processing. We could see the processed form of Sde2 from those mutated chimeras. This pointed out to some changes in the amino acid sequence in the Hub1 part of Hub1-$^{Leu}_{73}$Sde2-C chimera as now the DUBs are able to recognize it and generate the Sde2-C. So, we thought that with this hydroxylamine based strategy we could evolve a processing deficient protein to a processing efficient one. But surprisingly, when the sequencing results were analysed we found that

those processing efficient chimeras were WT Sde2 whose origin in those mutated strains is not explainable at this time.

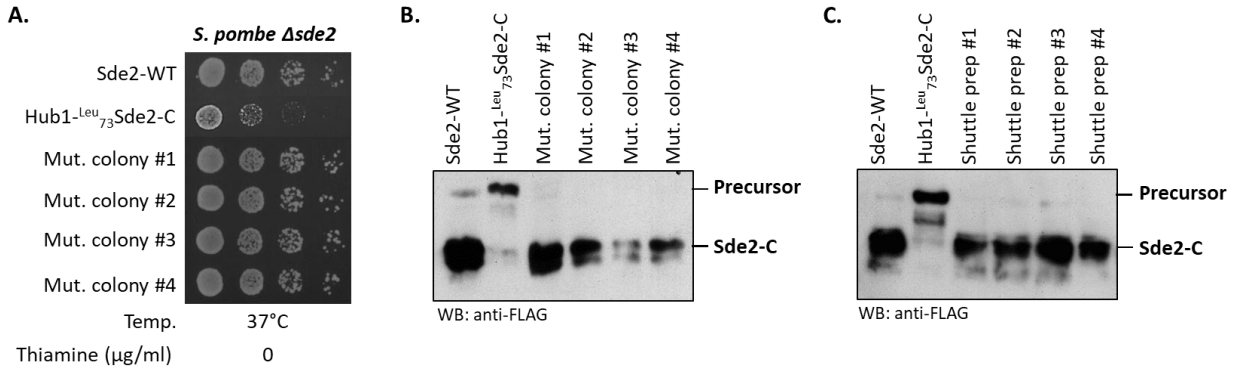


Figure 3.6.2: Screening from random mutagenized chimeras. A. Complementation of *S. pombe Δsde2* by strains having putative positively mutated Hub1-Sde2 chimeras after directed evolution. Fivefold serial dilution spot assay was done on indicated agar plates and was incubated at 37°C until growth appeared; B. Western blot analysis of those putative positive mutated chimeras; C. Plasmid was isolated from those mutated strains and retransformed into fresh cells. Western blot analysis to reconfirm the processing.

3.7 Function of $Sde2_{UBL}$

Ubiquitin and ubiquitin like proteins (ULM and UDP) are involved in virtually all cellular processes. Even though proteosomal degradation of proteins is the most discussed function there are other plenty of identified and yet 'to be explained' functions. The general scheme for ubiquitin or ubiquitin like proteins is that they synthesized as a precursor which upon co/post translational proteolytic processing (Grou et al, 2014) produces ubiquitin/UBL and the fusion tail protein. In case of ULMs, processing takes place after the typical diglycine motif and they function as post translational modifier. Ubiquitin is made from ubiquitin-ribosome and poly-ubiquitin precursor genes, whereas other ULMs like NEDD8, SUMO, ISG15 undergo processing to generate the ubiquitin fold domain proteins and short flanking tail. For some of the well characterized ULM, their function and substrate have been listed below-

UBLs	Functions	Substrate
Ubiquitin	Proteosomal degradation, cellular trafficking, DNA damage repair, ribosome biogenesis, immune system regulation.	Cytoplasmic and nuclear proteins, membrane receptors, histones, signaling proteins
NEDD8	Activation of Cullin E3 ligase	Cullin
SUMO	DNA repair, transcription, nuclear transport	Ran GTPase activating protein 1, TDG, activators and repressors of transcription
ISG15	Functions as cytokine, antiviral activity	Cytoplasmic and nuclear proteins
Atg12	Autophagy	Atg5

Table 1.1 Function and substrate of ubiquitin and some well-known UBLs

Following the same trend, Sde2 in that way is very unique; after processing of $Sde2_{UBL}$ ends with a characteristic diglycine motif and generates Sde2-C which is involved in splicing. No other ubiquitin like processing has been described to generate splicing factor before. Since no function of $Sde2_{UBL}$ has been assigned yet, it is very tempting to check whether, like ubiquitin, $Sde2_{UBL}$ makes conjugates to other proteins.

3.7.1 $Sde2_{UBL}$ as a post translational modifier

$Sde2_{UBL}$ has all the criteria to be a post translational modifier- first of all it harbors a characteristic ubiquitin like β-grasp fold and it contains diglycine motif which is required for isopeptide linkage for substrate conjugation. Ubiquitin like fold becomes very important as they mediate the recognition and interaction by ubiquitin binding proteins for downstream signaling processes (Husnjak et al, 2012). And the terminal glycine is needed to make isopeptide bond with the ε-amino group of acceptor lysine on the substrate proteins. With these set of informations in hand, we hypothesized $Sde2_{UBL}$ may get covalently conjugated to some proteins. In order to check that we took the simple minimalistic approach where we only expressed the $Sde2_{UBL}$ with N-terminal MYC tag in WT strains of *S. pombe*. We also

expressed the same construct in *Δubp5 Δubp15* strain since in deletion of those DUBs there should be an enrichment of the conjugates. The idea was that if $Sde2_{UBL}$ makes some sort of conjugates then simply by overexposure of the western blot one might be able to capture the conjugates. We could indeed see some higher molecular adducts of $Sde2_{UBL}$, but the nature of those adducts was not clear, whether they are SDS resistant non covalent or covalent higher molecular adducts. Previously, ubiquitin like protein Hub1 has been shown to form SDS resistant adducts in an ATP independent manner (Lüders et al, 2003); it is quite possible that $Sde2_{UBL}$ makes similar kind of adducts.

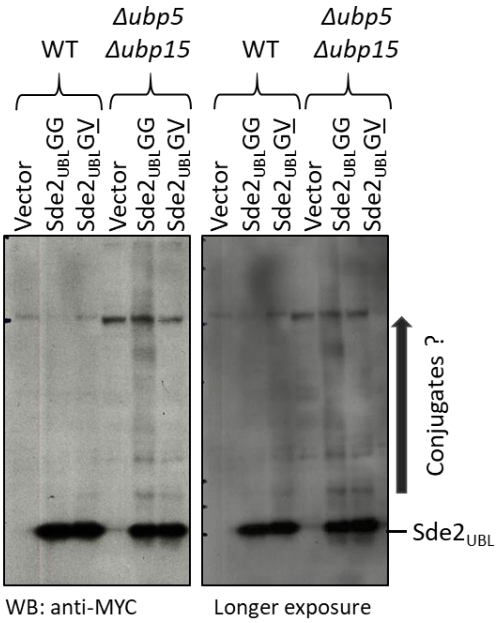


Figure 3.7.1. Possible function of $Sde2_{UBL}$ as a post translational modifier. $Sde2_{UBL}$ was transformed in WT and in *Δubp5 Δubp15* cells under constitutive promoter, western blot was performed to look at the higher molecular adducts in presence of SDS (courtesy Rakesh).

At this point it is not quite clear if those adducts have some physiological roles or not, one way to understand this better would be to take proteomics approach where we can identify if any cellular protein gets $Sde2_{UBL}$ conjugated and if there is any enrichment of those in *Δubp5 Δubp15* strain in comparison to WT strain.

3.7.2 Regulatory role of $Sde2_{UBL}$

Previously it has been shown that the precursor Sde2 association with the spliceosome is diminished (Thakran et al, 2017), which suggests that when processing is defective, the precursor of Sde2 does not end up going to the spliceosome to the full extent and hence splicing is affected. This could occur if $Sde2_{UBL}$ is inhibitory for Sde2-C incorporation in the spliceosome. To test whether this inhibitory role of $Sde2_{UBL}$ is Sde2 specific, we fused both processing efficient and deficient $Sde2_{UBL}$ to Hub1 at its N-term. The idea behind was to check whether $Sde2_{UBL}$-Hub1 chimera associates with the spliceosome. We hypothesized that Hub1 fusion with processing efficient $Sde2_{UBL}$ having the diglycine motif ($Sde2_{UBL}$GG-Hub1) would allow the Ubp5 and Ubp15 to cleave the $Sde2_{UBL}$ and release free Hub1 which could incorporate in the spliceosome; whereas in case of the processing defective $Sde2_{UBL}$ Hub1 chimera ($Sde2_{UBL}$GV-Hub1) DUBs would unable to cleave $Sde2_{UBL}$, resulting 'inhibitory' $Sde2_{UBL}$ to be tagged with Hub1. We took the *S. pombe* strain where Hub1 is deleted endogenously but maintained by a Uracil containing plasmid (Hub1 is essential in *S. pombe*) and transformed the $Sde2_{UBL}$GG/GV-Hub1 chimeras and looked for complementation. We could see that the processing efficient chimera was able to complement the defect when the uracil plasmid was shuffled out by 5-fluoroorotic acid (FOA), but surprisingly even the processing defective chimera complemented the growth defect. This suggests $Sde2_{UBL}$GV is not inhibitory for Hub1 association with the spliceosome contrary to Sde2-C.

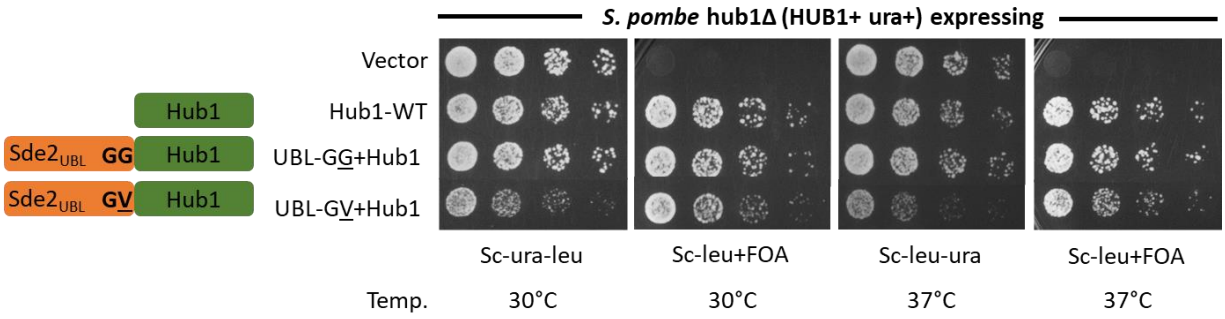


Figure 3.7.2: Regulatory role of $Sde2_{UBL}$. Complementation of *S. pombe* hub1Δ (HUB1+ ura^+) by Hub1 and $Sde2_{UBL}$-Hub1 chimera, constructs were expressed under constitutive promoter. Fivefold serial dilution spot assay was done on indicated agar plates and was incubated at 30°C and 37°C until growth appeared.

This difference could arise because Hub1 being a transient interactor of spliceosome does not require stable association with spliceosome or $Sde2_{UBL}$GV does not affect its function within the spliceosome. This could be due to the fact that the mode of function carried out by Sde2 and Hub1 in the spliceosome

are very different. Sde2 needs a free lysine at the C-terminus which is other masked by Sde2 $_{UBL}$ for proper splicing, whereas Hub1 modifies or associate with the spliceosome differently. So, it may be right to say at this point that in general Sde2$_{UBL}$ is not inhibitory for spliceosomal association, but it could be specific for Sde2.

Chapter 4

Discussion

In this study we have tried to find out two important aspects related to $Sde2_{UBL}$. First, how does $Sde2_{UBL}$ get recognized by two deubiquitinating enzymes Ubp5 and Ubp15 in *S. pombe*, what are the critical residues present on $Sde2_{UBL}$ which makes Sde2 a substrate of DUBs? And second, we tried to figure out possible function played by $Sde2_{UBL}$.

The dual specificity showed by Ubp5 and Ubp15 cannot be a random act of enzyme-substrate interaction, there must be a significant biological role which governs this whole process. Other than Sde2, no other splicing factor has been reported to show ubiquitin like processing by two canonical DUBs. In case of Sde2, generation of the Sde2-C is essential for intron specific pre-mRNA splicing. The low sequence similarity with ubiquitin might be the reason for $Sde2_{UBL}$ not to get processed by any other DUBs, but the fact that despite this Ubp5 and Ubp15 cleave Sde2 into functional form makes the process interesting. From multiple sequence alignment of Sde2 orthologs in other higher eukaryotes, it is apparent that only the GGKGG motif is absolutely conserved with moderate sequence homology in the C-term. This is mostly due to the functional conservation of Sde2-C which could be involved in RNA splicing. But extremely low sequence similarity in the UBL region may indicate presence of a non-functional regulatory domain, and hence no sequence conservation. Over the course of evolution $Sde2_{UBL}$ might lose its functionality and hence the conservation constrain to maintain it, still the N-terminal region manages to adopt the β-grasp fold like ubiquitin. This conservation of ubiquitin like fold might have got selected in evolution to be regulated by DUB activity. Presence of a conserved diglycine motif GGKGG hints towards the ubiquitin like processing in other orthologs as well.

Sde2 seems to be present only in intron rich organisms which might require special splicing factors to facilitate the removal of some weak introns. As the complexity of the organism has increased, the cellular proteome had evolved to cope up with it. DUBs need to evolve as the substrate proteins for ubiquitination has changed over time and gained more complexity. It is quite possible that along with Sde2, its processing DUBs have coevolved too; it can be speculated from the observation that Ubp15 of *S. cerevisiae* does not cleave *S. pombe* Sde2, similarly the mammalian ortholog of *S. pombe* Ubp15, Usp7 does not cleave *S. pombe* Sde2 as well (previously shown by Prashant). This suggests that as Sde2 has evolved, its processing DUB has retained the activity by changing some residues in its regulatory region

which is important for Sde2 recognition. Orthologs of Ubp5 and Ubp15 in other species have taken unique path in their evolutionary path in order to cleave their corresponding Sde2 ortholog. That is why only Ubp5 and Ubp15 are the processing enzymes for Sde2 at least in *S. pombe*, whereas some other protease may be gained that function in higher eukaryotes.

Now the issue of Sde2 having a ubiquitin like processing can be answered like this- that since Sde2 harbors a UBL with less sequence similarity to ubiquitin at its N-term followed by typical diglycine motif, any potential DUB could process it. This could produce Sde2-C constitutively without any regulation and might lead to nonspecific RNA splicing events. Whereas if only two specific DUBs are dedicated for this role to play along with their canonical DUB activity then Sde2 production can be under tight regulation. After processing Sde2-C associates with the spliceosome and it has been shown that ^{Lys}Sde2-C interacts with other splicing factor more efficiently. Generation of lysine can only be achieved when the UBL from Sde2 N-term is cleaved. Sde2$_{UBL,}$ thus clearly is important for generating the Sde2-C. But whether it has any other regulatory role function or not, it needs more detailed investigation. Nonetheless *S. pombe* Sde2 thus provide an appropriate platform to investigate ubiquitin like processing and its regulation, regulation of splicing and it also shows the connection between posttranslational modifications with RNA splicing.

Reference

1. Domains, U., Burroughs, A. M., Iyer, L. M., & Aravind, L. (2016). *Structure and Evolution of Ubiquitin and Chapter 2 Structure and Evolution of Ubiquitin and Ubiquitin-Related Domains*. https://doi.org/10.1007/978-1-61779-474-2
2. Kleiger, G., & Mayor, T. (2014). Perilous journey: A tour of the ubiquitin-proteasome system. *Trends in Cell Biology*, *24*(6), 352–359. https://doi.org/10.1016/j.tcb.2013.12.003
3. Herrmann, J., Lerman, L. O., & Lerman, A. (2007). Ubiquitin and ubiquitin-like proteins in protein regulation. *Circulation Research*, *100*(9), 1276–1291. https://doi.org/10.1161/01.RES.0000264500.11888.f0
4. Cappadocia, L., & Lima, C. D. (2018). Ubiquitin-like Protein Conjugation: Structures, Chemistry, and Mechanism. *Chemical Reviews*, *118*(3), 889–918. https://doi.org/10.1021/acs.chemrev.6b00737
5. Mishra, S. K., Ammon, T., Popowicz, G. M., Krajewski, M., Nagel, R. J., Ares, M., Jentsch, S. (2011). Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. *Nature*, *474*(7350), 173–180. https://doi.org/10.1038/nature10143
6. Sugioka-Sugiyama, R., & Sugiyama, T. (2011). Sde2: A novel nuclear protein essential for telomeric silencing and genomic stability in Schizosaccharomyces pombe. *Biochemical and Biophysical Research Communications*, *406*(3), 444–448. https://doi.org/10.1016/j.bbrc.2011.02.068
7. Thakran, P., Pandit, P. A., Datta, S., Kolathur, K. K., Pleiss, J. A., & Mishra, S. K. (2017). Sde2 is an intron-specific pre-mRNA splicing regulator activated by ubiquitin-like processing. *The EMBO Journal*, e201796751. https://doi.org/10.15252/embj.201796751
8. Jo, U., Cai, W., Wang, J., Kwon, Y., D'Andrea, A. D., & Kim, H. (2016). PCNA-Dependent Cleavage and Degradation of SDE2 Regulates Response to Replication Stress. *PLoS Genetics*, *12*(12), 1–26. https://doi.org/10.1371/journal.pgen.1006465
9. Xie, Y., Zhao, Y., Tan, H., Chen, Q., & Huang, J. (2017). An ubiquitin-like protein SDE2 negatively affects sucrose-induced anthocyanin biosynthesis in Arabidopsis. *Science Bulletin*, *62*(23), 1585–1592. https://doi.org/10.1016/j.scib.2017.11.014
10. Komander, D., Clague, M. J., & Urbé, S. (2009). Breaking the chains: Structure and function of the deubiquitinases. *Nature Reviews Molecular Cell Biology*, *10*(8), 550–563. https://doi.org/10.1038/nrm2731
11. Ronau, J. A., Beckmann, J. F., & Hochstrasser, M. (2016). Substrate specificity of the ubiquitin and Ubl proteases. *Cell Research*, *26*(4), 441–456. https://doi.org/10.1038/cr.2016.38
12. Bachmair, A., Finley, D., & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino-terminal residue. *Science*, *234*(4773), 179–186. https://doi.org/10.1126/science.3018930
13. Varshavsky, A. (1996). The N-end rule: functions, mysteries, uses. *Proceedings of the National Academy of Sciences*, *93*(22), 12142–12149. https://doi.org/10.1073/pnas.93.22.12142
14. Kouranti, I., Mclean, J. R., Feoktistova, A., Liang, P., Johnson, A. E., Roberts-Galbraith, R. H., & Gould, K. L. (2010). A global census of fission yeast deubiquitinating enzyme localization and interaction networks reveals distinct compartmentalization profiles and overlapping functions in endocytosis and polarity. *PLoS Biology*, *8*(9). https://doi.org/10.1371/journal.pbio.1000471
15. Nijman, S. M. B., Luna-Vargas, M. P. A., Velds, A., Brummelkamp, T. R., Dirac, A. M. G., Sixma, T. K., & Bernards, R. (2005). A genomic and functional inventory of deubiquitinating enzymes. *Cell*, *123*(5), 773–786. https://doi.org/10.1016/j.cell.2005.11.007

16. Komander, D., Reyes-Turcu, F., Licchesi, J. D. F., Odenwaelder, P., Wilkinson, K. D., & Barford, D. (2009). Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Reports*, *10*(5), 466–473. https://doi.org/10.1038/embor.2009.55
17. Gong, L., Kamitani, T., Millas, S., & Yeh, E. T. H. (2000). Identification of a Novel Isopeptidase with Dual Specificity for Ubiquitin-and NEDD8-conjugated Proteins*, *275*(19), 14212–14216. Retrieved from http://www.jbc.org/content/275/19/14212.full.pdf
18. Shin, Y. C., Chen, J. H., & Chang, S. C. (2017). The molecular determinants for distinguishing between ubiquitin and NEDD8 by USP2. *Scientific Reports*, *7*(1), 2–11. https://doi.org/10.1038/s41598-017-02322-x
19. Grou, C. P., Pinto, M. P., Mendes, A. V., Domingues, P., & Azevedo, J. E. (2015). The de novo synthesis of ubiquitin: Identification of deubiquitinases acting on ubiquitin precursors. *Scientific Reports*, *5*, 1–16. https://doi.org/10.1038/srep12836
20. Olvera, J., & Wool, I. G. (1993). The carboxyl extension of a ubiquitin-like protein is rat ribosomal protein S30. *Journal of Biological Chemistry*, *268*(24), 17967–17974.
21. Lacombe, T., García-Gómez, J. J., De La Cruz, J., Roser, D., Hurt, E., Linder, P., & Kressler, D. (2009). Linear ubiquitin fusion to Rps31 and its subsequent cleavage are required for the efficient production and functional integrity of 40S ribosomal subunits. *Molecular Microbiology*, *72*(1), 69–84. https://doi.org/10.1111/j.1365-2958.2009.06622.x
22. Pfiffner, J., Margalit, H., Habib, N., & Lin, M. F. (2011). Comparative Functional Genomics of the Fission Yeasts. *Science*, *332*(6032), 930–936. https://doi.org/10.1126/science.1203357.
23. Husnjak, K., & Dikic, I. (2012). Ubiquitin-Binding Proteins: Decoders of Ubiquitin-Mediated Cellular Functions. *Annual Review of Biochemistry*, *81*(1), 291–322. https://doi.org/10.1146/annurev-biochem-051810-094654
24. Lüders, J., Pyrowolakis, G., & Jentsch, S. (2003). The ubiquitin-like protein HUB1 forms SDS-resistant complexes with cellular proteins in the absence of ATP. *EMBO Reports*, *4*(12), 1169–1174. https://doi.org/10.1038/sj.embor.7400025
25. Archibald, J. M., Teh, E. M., & Keeling, P. J. (2003). Novel ubiquitin fusion proteins: Ribosomal protein P1 and actin. *Journal of Molecular Biology*, *328*(4), 771–778. https://doi.org/10.1016/S0022-2836(03)00374-7
26. Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, *9*, 1–8. https://doi.org/10.1186/1471-2105-9-40
27. Knop, M., Siegers, K., Pereira, G., Zachariae, W., Winsor, B., Nasmyth, K., & Schiebel, E. (1999). Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. *Yeast*, *15*(10B), 963–972. https://doi.org/10.1002/(SICI)1097-0061(199907)15:10B<963::AID-YEA399>3.0.CO

Appendix

Strain list

Strain	Relevant genotype
SP1	h- *ade6-M216, leu1, ura4-D18*
SP13	JY741 *hub1::aur1R pUR19-hub1+*
SP20	h+ *JY741 Δsde2:: Nat-NT2*

Plasmid list

Stock ID.	Plasmid
D052	pREP81x-3MYC–sde2–3FLAG
D057	pREP81x-3MYC–sde2(AAKGG)–3FLAG
D246	pREP81x-3MYC–Sde2(GAKGG)–3FLAG
D247	pREP81x-3MYC–Sde2(GVKGG)–3FLAG
D263	pENO-3MYC-SpSde2-3FLAG
D300	pREP81x-hub1GG–Leu_{73}sde2-C–3FLAG
D301	pREP81x-$_{1-22}$sde2-$_{19-73}$hub1-L_{73}sde2-C–3FLAG
D302	pENO-3MYC— $sde2_{UBL}$GG-hub1
D303	pENO-3MYC— $sde2_{UBL}$GV-hub1
D304	pENO-3MYC—$sde2_{UBL}$GG
D305	pENO-3MYC—$sde2_{UBL}$GV
D306	pENO-3MYC-hub1
D307	pREP81X-3MYC-hub1GG-KGGSde2-C—3FLAG
D308	pENO-3MYC-*S. japonicus* Sde2-3FLAG