Activation of a ubiquitin fold containing pre-mRNA splicing regulator Sde2 by deubiquitinating enzymes Ubp5 and Ubp15 in *Schizosaccharomyces pombe*

A Thesis

Submitted by

PRASHANT ARUN PANDIT

PH12112

For the award of the degree of

DOCTOR OF PHILOSOPHY

Department of Biological Sciences

Indian Institute of Science Education and Research (IISER) Mohali

Sector-81, Mohali-140306, Punjab, India

February 2019

Declaration

The work presented in this thesis entitled "Activation of a ubiquitin fold containing pre-mRNA splicing regulator Sde2 by deubiquitinating enzymes Ubp5 and Ubp15 in *Schizosaccharomyces pombe*" has been carried out by me under the supervision of Dr. Shravan Kumar Mishra at the Department of Biological Sciences, Indian Institute of Science Education and Research (IISER) Mohali.

This work has not been submitted in part or full for a degree, a diploma, or a fellowship to any 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 bona fide record of original work done by me and all sources listed within have been detailed in the references.

Date: Prashant Arun Pandit

Place: (Candidate)

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

Date: Dr. Shravan Kumar Mishra

Place: (Supervisor)

Acknowledgement

I express my gratitude to my Ph.D. supervisor Dr. Shravan Kumar Mishra, who is a great mentor and I am fortunate enough to get the chance to work with him. His cordial nature, guidance, scientific knowledge, experience and dedication made my work enjoyable and comprehensive. His excellent mentorship has immensely helped me in a better understanding of the work. He critically analysed the results and encouraged us to plan the experiments carefully with utmost perfection. He emphasized on being focused, organized, and disciplined in the field of research. I am thankful for his time and immense efforts for the successful completion of the project.

I express my sincere thanks to my doctoral committee members, Prof. Anand Kumar Bachhawat and Dr. Kausik Chattopadhyay for their time, support, and valuable suggestions. Their critical comments and guidance have always been very helpful for the timely progress of my doctoral study and enhanced the quality of my work.

I would like to thank Prof. N. Sathyamurthy, former Director, IISER Mohali, for giving me the opportunity to work at this premier research institute. I am also thankful to Prof. Debi Prasad Sarkar, the Director, IISER Mohali for kindly permitting me to use the excellent infrastructure for carrying out my research work.

I profoundly thank our collaborator Dr. Jeffrey A. Pleiss at Cornell University, Ithaca, USA for sharing the splicing-sensitive microarrays design, helping us to perform the experiment and for the analysis of data. Special thanks to Prof. Stefan Jentsch at the Max Planck Institute of Biochemistry, Martinsried, Germany for his inputs, support and timely review of the progress of the work. I acknowledge Mrs. Ulla Cramer from Prof. Stefan Jentsch laboratory at the Max Planck Institute of Biochemistry, Martinsried, Germany and NBRP/YGRC, Japan for materials and strains used in this study.

Special thanks to Poonam, Sumanjit and Kiran for their immense contribution in the successful completion of the project. I am extremely thankful to all my lab members Poulami, Rakesh, Karan, Praver, Anupa, Manu, Ankita, Sabbir for their incredible support, fruitful scientific discussions and maintaining healthy environment in the lab. I thank Mr. Baidnath Mandal (Vidhya) for his efforts and support in the lab work.

I express my immense gratitude to all faculties and students of Department of Biological Sciences, IISER Mohali for providing me access to departmental facilities, for the exciting discussions and encouraging interactions during seminars and presentations.

I feel myself lucky to be a part of PhD Aug 2012 batch and I sincerely thank all my batch mates at IISER for making tenure of my PhD cheerful, adventurous and enjoyable. I acknowledge Bhupinder, Devashish, Krishna, Prince, Shashank, Shiv, Soumitro and Nagesh for the scientific and non-scientific discussions and their support.

I owe my deepest gratitude to my family for their blessings, sacrifices, and faith. Their support has always strengthened and inspired me to achieve my dreams.

I am very thankful to all non-teaching staff members, administration, library staff, store and purchase sections, accounts sections, hostel caretakers, mess peoples, cleaning staff and entire IISER Mohali community for their help and cooperation.

I would like to thank University Grants Commission, Government of India, for the fellowship during my PhD tenure, IISER Mohali, Centre for Protein Science, Design and Engineering (CPSDE) at IISER Mohali, Ministry of Human Resource and Development (MHRD), DST, Government of India, Max Planck Society, Germany, for providing financial support and EMBO for providing international travel fellowship award.

Table of contents

Contents	Page no.

List of Tables

List of Figures

Contents **Page no.**

Abbreviations

aa	amino acids
AS	Alternative Splicing
Atg	Autophagy-related gene
ATM	Ataxia Telangiectasia Mutated
ATP	Adenosine Triphosphate
CP	Core Particle
CRLs	Cullin-RING E3 Ligases
CSN	COP9 signalosome
Cys	Cysteine
DUB	Deubiquitinating enzyme
ER	Endoplasmic Reticulum
ESE	Exonic Splicing Enhancer
ESS	Exonic Splicing Silencer
ESCRT	Endosomal Sorting Complex Required for Transport
FAT10	HLA-F adjacent transcript 10
Gly	Glycine
GMP1	GAP Modifying Protein1
GPCR	G-Protein Coupled Receptor
Hub1	Homologous to ubiquitin
ISS	Intronic Splicing Silencer
IFN	Interferon
JAMM	JAB1/MPN/MOV34
Lys	Lysine
MATH	Meprin and TRAF Homology
Met	Methionine
MIU	Motif Interacting with Ubiquitin
MPT	Molybdopterin
Myb	Myoblastosis
NEDD8	Neural precursor cell Expressed Developmentally Downregulated
NEDP1	Nedd8-specific Protease1
NPC	Nuclear Pore Complex
OTU	Ovarian Tumor Protease

PCNA	Proliferating Cell Nuclear Antigen
PIAS1	Protein Inhibitor of Activated STAT-1
PML	Promyelocytic Leukemia protein
PPi	Pyrophosphate
PPT	Polypyrimidine tract
Pre-mRNA	Precursor messenger RNA
Prp	Pre-mRNA Processing protein
PTMs	Post-Translational Modifications
RBP	RNA Binding Protein
RNP	Ribonucleoprotein
ROS	Reactive Oxygen Species
RP	Regulatory Particle
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
RT-qPCR	Real-Time quantitative PCR
SAGA	Spt-Ada-Gcn5-Acetyltransferase complex
SENP	sentrin/SUMO-specific protease
SMN	Survival Motor Neuron
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
SUMO	Small Ubiquitin-like Modifier
TLS	Translesion Synthesis
TNF	Tumour Necrosis Factor
Ub	Ubiquitin
UBL	Ubiquitin-like protein
UBA	Ubiquitin-Associated domains
UCH	Ubiquitin Carboxy-terminal Hydrolase
UFM1	Ubiquitin Fold Modifier
UIM	Ubiquitin-Interacting Motifs
ULP	UBL-specific protease
URM1	Ubiquitin-Related Modifier
USP	Ubiquitin-Specific Protease

Synopsis

Introduction

The pre-mRNA splicing is one of the fundamental processes in the cell which governs the expression of genetic information. A multimegadalton ribonucleoprotein complex termed spliceosome removes the non-coding introns and joins the exons to generate translatable mRNA. Splicing defects thereby affect protein expression and can cause cellular dysfunction or cell death. Pre-mRNA splicing is regulated through protein/RNA interactions and post-translational modifications of spliceosomal components including ubiquitin (Ub) and ubiquitin-like modifiers (UBLs) [1].

The covalent or non-covalent modification of protein substrates by Ub and UBLs play a key regulatory role in almost all cellular functions [2]. Although their amino acid sequences vary, they share a β-grasp fold and thus show structural homology. They are often synthesized as precursors which get processed and activated by deubiquitinating enzymes (DUBs) and UBL-specific proteases respectively. These enzymes are diverse in their domain architectures and highly substrate specific [3].

The UBL Hub1 is known to regulate alternative splicing in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* and humans [4,5]. The protein is highly conserved and contains ubiquitin fold, but shows very low sequence similarity to ubiquitin. *S. pombe* Hub1 genetically interacted with multiple splicing mutants where Sde2 appeared as one of the candidate.

The predicted structure of Sde2 shows a ubiquitin-fold domain ($Sde2_{UBL}$) at the amino terminus followed by an invariant GGKGG motif and a helical carboxy-terminal domain (Sde2-C). *S. pombe* strains lacking Sde2 gene showed growth defects, temperature sensitivity, defects in telomeric silencing and genomic stability [6], and recently the protein was shown to co-purify with splicing factors [7,8]. The presence of a conserved di-glycine motif after $Sde2_{UBL}$ and genetic interaction of Sde2 with Hub1 encouraged us to study the role Sde2 in *S. pombe*.

Objectives of the study

1. The predicted Sde2 structure has a ubiquitin-fold domain and a conserved di-glycine motif. We asked whether Sde2 is cleaved like ubiquitin and whether the processing is required for its function.

2. If Sde2 indeed got processed, then what are its processing enzymes and what is the role of $Sde2_{UBL}$?
3. Sde2 associated with the spliceosome. We wanted to know its mode of spliceosomal association and function.

Outcomes of the study

USP-domain proteases Ubp5 and Ubp15 process S. pombe Sde2 precursor

We expressed an epitope-tagged version of Sde2 with *S. pombe* Δ*sde2* strain to detect Sde2 protein by immunoblot assays. Sde2 full-length protein got cleaved into $Sde2_{UBL}$ and Sde2-C. The processing is presumed to happen at GG~KGG site where alanine substitutions of the first GG residues abolished the processing. To find out the processing enzymes, we expressed Sde2 in protease deletion strains available from Bioneer *S. pombe* deletion library. Among protease deletion mutants, Δ*ubp15* was the only mutant showing partial accumulation of Sde2 precursor. Ubp15 is a ubiquitin specific protease (USP) domain containing DUB which is known to cleave ubiquitin. It also harbours a MATH (Meprin and TRAF homology) domain and C-terminal domain containing five UBL-like repeats. Since Ubp5 is the paralog of Ubp15 in *S. pombe*, Δ*ubp5* Δ*ubp15* double deletion led to complete accumulation of the precursor. Co-expression of *S. pombe* Sde2 with either Ubp5 or Ubp15 in *E. coli* confirmed the processing of Sde2 by these DUBs.

Processing of Sde2 facilitates its incorporation into the spliceosome

Co-immunoprecipitation and mass spectrometry experiment (performed by my colleagues) revealed that Sde2 co-purified with spliceosomal proteins. Using chromosomally tagged strains of Sde2 gene harbouring mutations at the GGKGG site, we monitored growth phenotypes and spliceosomal incorporation of the mutants. With co-IP experiments of splicing factors, we observed a reduced spliceosomal association of the processing defective AAKGG mutant. We also observed that *S. pombe* Δ*ubp5* Δ*ubp15* strain, where Sde2 processing was completely abolished, showed splicing defects like Δ*sde2* strain. Thus, Sde2 processing by Ubp5 and Ubp15 is essential for its function, and the UBL fold appears to have an inhibitory role for Sde2's incorporation into the spliceosome.

$Sde2_{UBL}$ supports optimal expression of Sde2-C protein

In ubiquitin-ribosomal fusion proteins, ubiquitin has a co-translational chaperone-like function for the ribosomal proteins [9]. To monitor whether $Sde2_{UBL}$ has similar function, we

removed the $Sde2_{UBL}$ region from its chromosomal locus to study its effect on Sde2-C protein. The mutant *S. pombe* strain with *de novo* synthesized Met-Sde2-C showed growth defect and lower protein levels. However, the spliceosomal incorporation of Met-Sde2-C lacking $Sde2_{UBL}$ remained unaffected. Sde2-C was optimally expressed when processed from the precursor fused to $Sde2_{UBL}$. The ubiquitin fold of Sde2 therefore may play a similar role like ubiquitin to stabilize Sde2-C.

Sde2-C with lysine at the N-terminus facilitates recruitment of Cay1 to the spliceosome

In a comparative mass spectrometry analysis of Prp19 and Cdc5 co-purified spliceosomes (performed by my colleagues), the spliceosomal association of Cay1 was reduced in the absence of Sde2. Since lysine to methionine mutants of Sde2 (GGKGG to *GGMGG*) showed growth defects, but no processing defects, we questioned whether the lysine at N-terminus of Sde2-C has any role in the spliceosomal association of Cay1. For this we made additional chromosomal mutants of lysine to arginine (*GGRGG*) and threonine (*GGTGG*). Above mutants resulted in diminished associations with Cay1 and showed Δ*sde2*-like splicing defects. These findings explained the role of $Sde2_{UBL}$ in generating a functional form of Sde2-C starting with the conserved lysine.

MATH and USP domain mutations of Ubp5 and Ubp15 affect processing of Sde2 in vitro

The DUBs Ubp5 and Ubp15 are dual specific proteases which act on both ubiquitin and Sde2. The human homolog of Ubp5 and Ubp15, i.e., USP7, did not process *S. pombe* Sde2 or *Hs* Sde2, but USP7 cleaved ubiquitin fusions in a recombinant system. Similar to Ubp5 and Ubp15, *Hs*USP7 also harbours a MATH domain followed by a USP catalytic domain and a carboxy-terminal domain containing five UBLs. To know the region responsible for the processing activity of Ubp5 and Ubp15 on Sde2, we made truncations in Ubp5 or Ubp15 and domain-swapped chimeras with USP7. Ubp5 truncation mutation without the MATH domain or with a truncated USP domain resulted in diminished activity on Sde2, whereas a Ubp15 truncation mutation without the UBL domains could still cleave Sde2 *in vitro*. Hence, MATH and USP domains of Ubp5 or Ubp15 are necessary for Sde2 recognition and processing.

Conclusion

Through my doctoral study, I have discovered two dual-specific DUBs Ubp5 and Ubp15 which activate a specialized pre-mRNA splicing regulator Sde2. The activation of Sde2 is unique since it is possible only by enzymes that cleave the precursor after 'GG' motif. Intriguingly, Sde2 cleaving enzymes are DUBs that recognize the two substrates which are

less than 20% identical to each other. We thereby have revealed a remarkable mechanism of splicing regulation where ubiquitin-like processing has been utilized to generate a specific splicing regulator.

References

1 Chanarat, S. and Mishra, S.K. (2018) Emerging Roles of Ubiquitin-like Proteins in Pre-mRNA Splicing. *Trends Biochem. Sci.* 43, 896–907

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

3 Mevissen, T.E.T. and Komander, D. (2017) Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* 86, 159–192

4 Mishra, S.K. *et al.* (2011) Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. *Nature* 474, 173–180

5 Ammon, T. *et al.* (2014) The conserved ubiquitin-like protein Hub 1 plays a critical role in splicing in human cells. *J. Mol. Cell Biol.* 6, 312–323

6 Sugioka-Sugiyama, R. and Sugiyama, T. (2011) Sde2: A novel nuclear protein essential for telomeric silencing and genomic stability in Schizosaccharomyces pombe. *Biochem. Biophys. Res. Commun.* 406, 444–8

7 Chen, W. *et al.* (2014) Endogenous U2·U5·U6 snRNA complexes in S. pombe are intron lariat spliceosomes. *RNA* 20, 308–20

8 Bayne, E.H. *et al.* (2014) A systematic genetic screen identifies new factors influencing centromeric heterochromatin integrity in fission yeast. *Genome Biol.* 15, 481

9 Lacombe, T. *et al.* (2009) Linear ubiquitin fusion to Rps31 and its subsequent cleavage are required for the efficient production and functional integrity of 40S ribosomal subunits. *Mol. Microbiol.* 72, 69–84

10 Clague, M.J. and Urbé, S. (2010) Ubiquitin: Same molecule, different degradation pathways. *Cell* 143, 682–685

Introduction

1.1 Ubiquitin system

Eukaryotic cells have several intracellular proteolytic pathways for degrading misfolded or denatured proteins and normal proteins whose level must be regulated for cellular homeostasis. There are three major protein degradation pathways used by cells, namely the lysosome, the proteasome, and the autophagosome. The ubiquitin (76 amino acid protein) is a conventional regulator of these pathways, where chemical modification of a target protein by addition of ubiquitin decides the fate of degradation [1]. The proteasomal degradation of target protein takes place in two successive steps: (i) the covalent conjugation of multiple ubiquitin molecules to the substrate and (ii) the degradation of the polyubiquitinated protein by 26S proteasome. Besides, the processes like cell cycle, transcription, cell proliferation, DNA replication, development, and splicing are controlled by degradative and non-degradative ubiquitin modifications. The regulatory network mediated through ubiquitin modifications and associated machinery is collectively referred to as 'ubiquitin system' [2].

1.1.1 Ubiquitin

In *Schizosaccharomyces pombe*, there are five ubiquitin coding genes namely ubi1, uep1/ubi2, ubi3, ubi4, and ubi5. The ubi1 and uep1 genes encode for a precursor of ubiquitin fused to ribosomal 60S subunit L40e. The ubi3 and ubi5 genes encode for precursors of ubiquitin fused to ribosomal 40S subunit S27a and S27b respectively. The ubi4 gene is expressed as precursor linear fusion of head to tail repeats of 5 ubiquitin molecules. It is reported in *Saccharomyces cerevisiae* that the N-terminal ubiquitin moiety acts as a chaperone to facilitate correct folding and efficient synthesis of Rps31, a subunit of the 40S ribosome [3]. Ubiquitin is a highly conserved 76-amino-acid residue protein which gets covalently conjugated to target protein. It is a highly stable protein which adopts a compact β-grasp fold structure with a flexible C-terminal tail [4]. There are only three amino acid differences between yeast and human ubiquitin sequence.

The conjugation of ubiquitin to target substrate requires three enzymes; the ubiquitin-activating enzyme E1, which activates the C-terminal Gly residue of ubiquitin in an ATP dependent manner. This reaction involves the formation of ubiquitin adenylate intermediate and release of PPi, followed by the binding of ubiquitin to a Cys residue of E1 through

thioester linkage and release of AMP. Activated ubiquitin molecule is transferred to ubiquitin-conjugating enzyme E2 by forming a thioester linkage with active site Cys residue. In the third and final step, ubiquitin ligase E3 catalyzes isopeptide bond formation between ubiquitin C-terminal Glycine and ε-amino group of Lysine of substrate protein [5]. The ubiquitin-ribosomal fusion proteins have evolved because the ribosomal counterparts are small charged proteins which lack sufficient hydrophobic surface for chaperone binding. Therefore ubiquitin fusion acts as a chaperone and prevents their premature degradation [6]. Ubiquitin has seven lysine residues, i.e., Lys 6, 11, 27, 29, 33, 48, 63, and N-term methionine residue which can participate in the substrate modification. Although polyubiquitin chains formed via Lys^{6}, Lys^{11}, Lys^{29}, and Lys^{63} are detected in cell-free systems and/or in whole cells [7–9], the Lys^{48}-linked polyubiquitin chains represent the predominant signal for targeting substrates to the 26S proteasome. Substrates with four or more linked ubiquitin molecules are often targeted to the 26S proteasome for degradation. Also, the monoubiquitination of substrates or its polyubiquitination through a range of lysine residues of ubiquitin can alter the localization or activity of the target protein. The factors that target a substrate to a particular pathway include chain length, linkage type or the lysine residue of ubiquitin involved in chain formation. These modifications of target proteins help in their association with receptors or mediate protein assembly (Fig.1.1A) [10].

1.1.2 The 26S Proteasome

The 26S proteasome is a large (~2.5 MDa) multi-catalytic ATP-dependent protease complex, and it serves as degrading machine of the ubiquitin-proteasome system [11]. It comprises two distinct subcomplexes, a 20S core particle (CP) and a 19S regulatory particle (RP) as shown in figure 1.1B [12]. The proteasome forms a barrel shape structure with four axially stacked heteroheptameric rings. The 20S CP is made up of two outer α rings (α1- α7 subunits) and two inner β rings (β1- β7 subunits). The α-rings form a pore which works as a gate for regulated entry of ubiquitinated substrates into the proteasome. The subunits of β rings have active sites with different proteolytic specificities: β1 with the peptidyl-glutamyl-hydrolyzing or caspase-like, β2 with the trypsin-like, and β7 with the chymotrypsin-like activity; therefore each mature proteasome have six proteolytic sites [13,14].

The 19S RP subunit has a regulatory role and takes care of identification, binding, deubiquitination, unfolding, and translocation of the substrate to the core particle. The “base” of 19S subunit contains six regulatory particles i.e., AAA ATPase (Rpt1-Rpt6) and four non-

ATPase (Rpn1, Rpn2, Rpn10, and Rpn13) subunits [12] and the "lid" of 19S contains nine different Rpn subunits (Rpn3, Rpn5-9, Rpn11, Rpn12 and Rpn15 (Dss1/Sem1)).

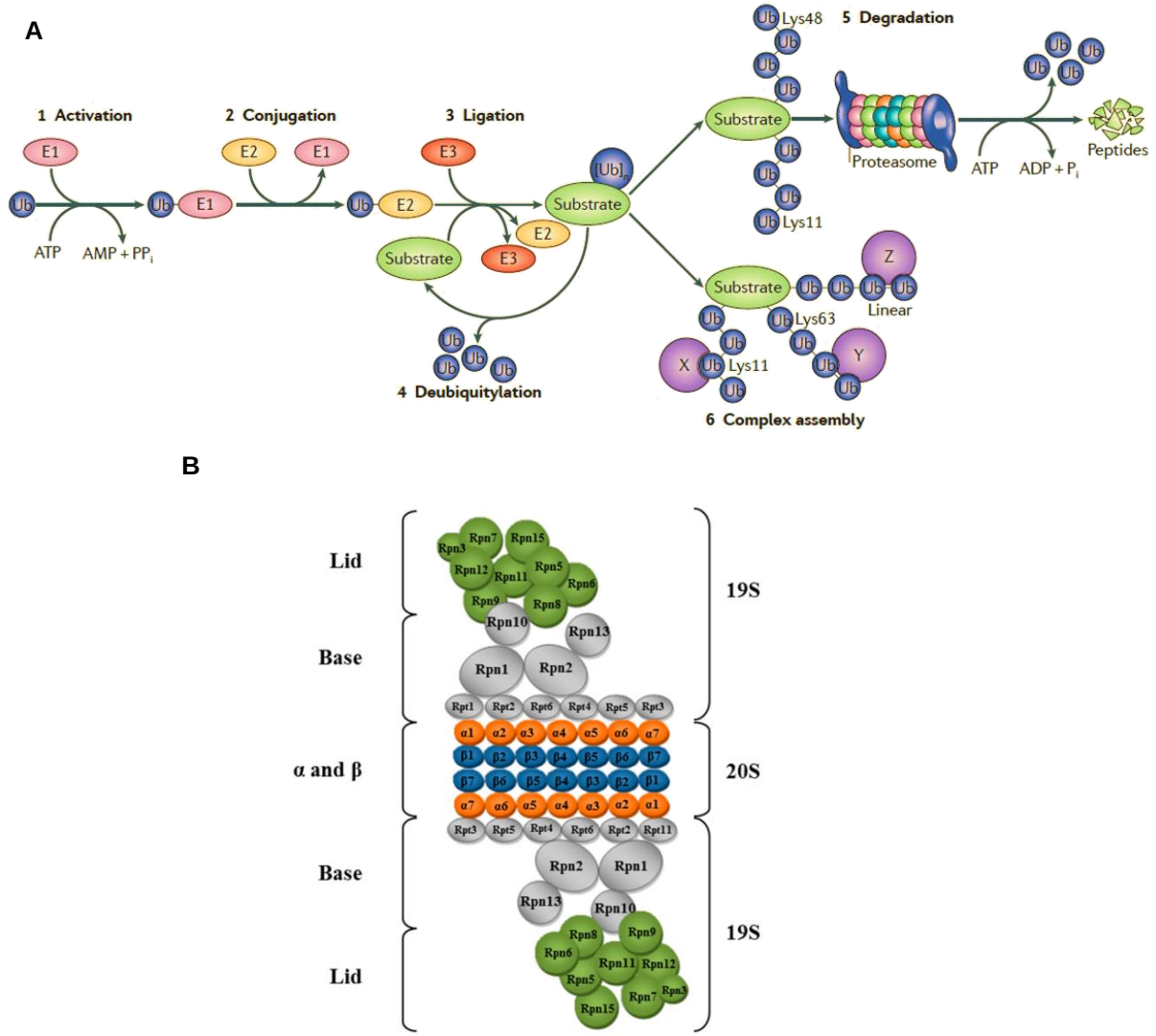


Figure 1.1 - The enzymes and reactions of the ubiquitin system.

A) The ubiquitylation and degradation of substrate proteins is achieved by a series of reactions mediated by the enzymes of the ubiquitin-proteasome system (UPS). In the activation reaction, ubiquitin is transferred to an E1 enzyme in an ATP-dependent manner (step **1**). Then to E2 enzyme in the conjugating reaction (step **2**). The E2 enzyme then carries the ubiquitin to the E3 enzyme, the ubiquitin ligase (step **3**). This process of repeated ubiquitin ligation with Lys of the ubiquitin itself leads to the formation of a polyubiquitin chain on the target protein. Deubiquitylating enzymes may reverse substrate protein ubiquitylation (step **4**). Lys11- and Lys48-linked polyubiquitin chains often target substrate proteins for proteasomal degradation (step **5**). Conversely, linear, Lys63- and Lys11-linked chains promote the assembly of signaling complexes (step **6**). X, Y, and Z indicate ubiquitin-binding proteins. Pi, inorganic phosphate; PPi, inorganic diphosphate; Ub, ubiquitin. The figure is adapted from [10].

B) **Scheme of the 26S proteasome.** Proteins that make up the base and lid of the 19S regulatory subunit are shown. The cylindrical portion of the 20S catalytic subunit is shown in an open conformation, showing the arrangement of α and β proteins identified in orange and blue, respectively. The figure is adapted from [12].

The lid carries out deubiquitination of incoming substrates with the help of deubiquitinating enzymes (DUBs) Rpn11, Uch37 and Ubp6/Usp14. The 19S and 20S sub-complexes are held together using high energy nucleotides, i.e., through binding of ATP with walker motifs of AAA ATPases of 19S subunit. The 19S binding results in conformational relief of the occlusion that blocks 20S proteasome pores. This phenomena opens the "gate" to the catalytic chamber by coordinating the timed separation and proper movement of the α-ring N-termini of 20S CP [15,16]. The opening of "gate" to the catalytic chamber, substrate processing, and translocation into CP for degradation are energy-dependent processes [11].

1.2 Deubiquitinating enzymes

The deubiquitylases (DUBs) are enzymes that mainly function in reversal of ubiquitin modifications and activation of ubiquitin precursors (Fig.1.2). The number of DUBs varies in different organisms; for example, around ~20 DUBs exists in *S. cerevisiae* or *S. pombe* whereas up to ~100 DUBs are found in humans [17]. Most of the DUB activity is cryptic in the sense that, the energy of associating with the substrate or a scaffolding protein is required to achieve the catalytically competent conformation. This means that unless the specific substrate is bound to these enzymes, the catalytically active conformation is not reached. This property helps in preventing cleavage of inappropriate substrates. DUBs are modular, i.e., in addition to the catalytic domain, they may contain accessory ubiquitin-binding domains or various protein-protein interaction domains. These modules contribute to the binding and recognition of substrate or help in the assembly of multiprotein complexes that localize DUBs to the substrate site. DUBs require these localization and substrate specificity determinants to function physiologically. The association of DUBs with substrate adapters, scaffolds, and inhibitors are regulatory interactions driving specificity [18].

The catalytic activity of DUBs is used to maintain free ubiquitin levels, rescue proteins from ubiquitin-mediated degradation, and control the dynamics of ubiquitin-mediated signaling pathways. Depending upon the complexity of ubiquitin modification involving either of the seven lysines or N-terminal methionine, the DUBs show specificities for deubiquitination, e.g., exo activity for removing ubiquitin at the distal end, endo activity for breaking the internal iso-peptide bond in polyubiquitin chain. The complexity further increases with heterotypic chains e.g., branched ubiquitin chains or mixed polyubiquitin chains involving UBLs like SUMO, NEDD8 and phosphorylated ubiquitin [19].

The DUBs are subdivided into six families based on the architecture of their catalytic domains. This includes four families of papain-like cysteine proteases: ubiquitin specific proteases (USPs), ubiquitin carboxy-terminal hydrolases (UCHs), ovarian tumor proteases (OTUs), Josephin domain, and a distinct JAB1/MPN/MOV34 (JAMM) domain family. Recently, the sixth member is discovered called as the motif interacting with ubiquitin (MIU)-containing novel DUB family (MINDYs) [20]. A separate class of proteases termed as ubiquitin-like protein specific proteases (ULPs) which act on ubiquitin-like modifiers such as SUMO [SENP (sentrin/SUMO-specific protease)] or NEDD8 [NEDP1 (NEDD8-specific protease 1)] [21,22].

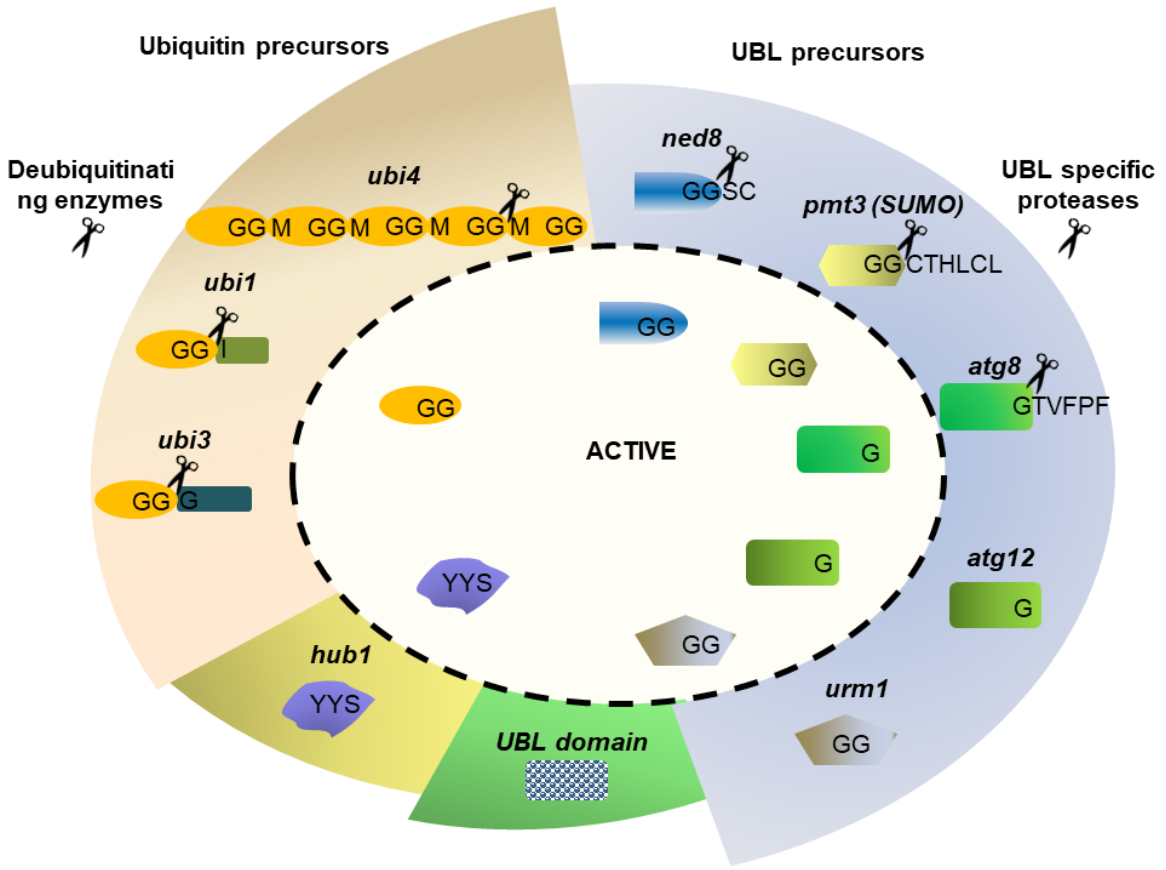


Figure 1.2 - Ubiquitin and UBL precursors in *S. pombe*
The Ub/UBLs are synthesized with C-terminal extension after the di-glycine motif. The processing site for Ubiquitin or UBL activation is indicated by scissor. Diglycine motif is absent in Hub1. The UBL domain-containing proteins also exists which share a β-grasp fold.

1.2.1 Ubiquitin-specific protease family

USP family DUBs are the largest and most diverse family of deubiquitinases. In yeast, there are 16 USP family DUBs, and humans are predicted to have over 50 members of this family [23]. Their length varies from 350 and 3400 amino acids (aa) with an averages length of ~1000 aa and possesses a core catalytic domain of ~350 aa. However, the length of the catalytic domain may vary from 450-800 amino acids. The USP domain consists of three subdomains, i.e., Palm, Thumb, and Fingers sub-domain together resembling a right hand like structure [24]. The catalytic center lies at the interface between Palm and Thumb, while the Fingers domain holds the distal ubiquitin. The globular portion of ubiquitin interacts with finger domain, and the C-terminal tail of ubiquitin extends into the cleft between thumb and palm sub-domains of the catalytic core. An exception to this fact is the CYLD protease which lacks finger domain. USP domain undergoes a dramatic conformational change upon ubiquitin binding. For example, the catalytic Cys residue in USP7 shifts upon ubiquitin binding, from a catalytically unproductive to an active position and interacts with the catalytic His residue [24,25].

The catalytic USP domain can be subdivided into six sequence boxes, where boundaries between the boxes are the point of large insertions occurred during evolution. These insertion points contain protein interacting domains, e.g., B-box in CYLD, MYND domain in USP19, etc. [26]. Several USP DUBs harbour additional ubiquitin-binding domains (UBDs) such as zinc-finger ubiquitin-specific protease (ZnF UBP) domain [17]. Among other domains are ubiquitin-interacting motifs (UIM), ubiquitin-associated (UBA) domains, and ubiquitin-like domains (UBL) which facilitate substrate interaction or subcellular localization or have an auto-regulatory role for enzyme activity. For example, UBA domains in USP5, UIM motifs in USP37, and UBL domains in USP7 [17,18,27].

1.2.2 Ubiquitin C-terminal hydrolase family

Humans have four members of the UCH family, and only one member is found in *S. cerevisiae* [18]. Their catalytic core is of ~260 amino acids as exemplified from structural studies of UCHL-1 and -3. UCHL1/3 mainly target ubiquitin precursor molecules with a short amino acid or peptide extension at the C-terminus and ubiquitin trapped in thiolester intermediates from ubiquitin conjugation pathways. UCHL-1 & -3 have roles in brain function [28], and a mutant of the former is associated with familial Parkinson's disease [29]. A third member UCHL5 has ~100 amino acids extension, and it interacts with Rpn11 in the proteasome. It functions in the recycling of ubiquitin chains from proteasome substrates

[30,31]. Fourth UCH family member BAP1 (BRCA1 associated protein-1) have an additional 500 amino acids and a nuclear localization sequence. It is a tumor suppressor and interacts with the BRCA1/BARD1 E3 ubiquitin ligase involved in DNA repair [32,33].

The catalytic residues of UCH family enzymes undergo a conformational change upon ubiquitin binding which generally exists in a non-productive conformation. The most striking feature of the UCH domain is a large crossover loop formed upon ubiquitin binding. The ubiquitin C-terminus has to extend through this loop to reach active site. This unique property of the UCH domain restricts the folded ubiquitinated substrates or ubiquitin polymers from accessing the active site. Hence they show negligible activity towards polyubiquitin chains with varied linkage types. UCHL5 probably attains a different conformation induced by proteasome for cleaving polyubiquitin chains [17,18,27].

1.2.3 Ovarian tumor domain family (OTU)

OTU domain family shows homology with ovarian tumor gene in fruit flies required for the development of ovaries. There are 15 OTU DUBs present in humans, and only 2 DUBs from OTU family are encoded in yeast [18]. The length of the OTU DUBs range from 230 to 1222 aa and contain additional domains like UBA, UIM, and UBL as mentioned before. Catalytic core is comprised of 150-200 residues. However, subclasses of OTU family enzymes like A20, Cezanne1/2, TRABID, and VCIP135 have an extended catalytic core of ~360 residues [27].

Similar to USP domain proteases, the distal ubiquitin binding site of OTU catalytic domain undergoes disordered to ordered transition upon ubiquitin binding. In case of OTUB1, the active site holds an unproductive state and requires conformational changes before activation. Such inactive state of catalytic core is found in cysteine proteases type DUBs which prevent the active site cysteine from oxidative stress. OTU domain proteases show marked chain linkage specificity like TRABID and DUBA (Lys63-specific), OTUB1 (Lys48-specific), etc. OTU family enzymes are involved in cell signaling such as NFκB signaling (A20, Cezanne1/2), Wnt signaling (TRABID) and IRF3 signaling (OTUD5) [18,27].

1.2.4 Josephin domain DUBs

Ataxin-3 is a prominent member of four human Josephin family proteins. Ataxin-3 is implicated in a neurodegenerative disorder called spinocerebellar ataxia type 3. A unique

feature of the Josephin domain is a large helical lever that restricts the access to the active site in the absence of ubiquitin. The NMR studies have revealed that ubiquitin binding stabilizes an active conformation of Ataxin-3. Interestingly, Ataxin-3 catalytic activity is activated by ubiquitination of the Josephin domain itself by an unknown E3 ligase. Speculations say that ubiquitination may stabilize an open helical lever confirmation. Ataxin-3 has three UIM motifs among which the two proximal UIMs interact with Lys48 linked ubiquitin chains. However, Ataxin-3 may edit Lys63 linked mix ubiquitin chains. Targets of Ataxin-3 and physiological role of other members are unclear [27].

1.2.5 JAB1/MPN/Mov34 domain family

JAMM/MPN+ family proteases are metalloenzymes where zinc ions coordinate with invariant His, Asp, and Ser residues for catalysis or may help in substrate recognition. [17][34]. There are eight human DUBs with JAMM domain, and they often operate as a part of the multi-subunit protein complex. These DUBs function in multiple pathways, for example, POH1 recycles ubiquitin chains in proteasome [35], while AMSH-LP works in vesicular trafficking and is part of ESCRT machinery [36,37]. BRCC1 is present in two DNA repair complexes, the BRISC complex and the BRCA1 A complex [27]. A COP9 signalosome component CSN5 is a deneddylating enzyme that removes Nedd8 modification from Cullin E3 ligases [38]. PRPF8 is a splicing factor harboring a JAMM/MPN+ domain with an impaired metal binding site. Hence it may have lost its enzymatic activity [39].

Most of the JAMM/MPN+ family proteases cleave lys63 linked ubiquitin chains and some members (AMSH, AMSH-LP, BRCC3) explicitly target lys63 linked chains. The crystal structure of AMSH-LP with lys63 linked diubiquitin gives insights into the mechanism of specificity. An extended conformation of lys63 linkage is stretched maximally by AMSH-LP to reach the active site. JAMM domain of AMSH-LP interacts with both the distal ubiquitin and the tripeptide sequence Gln62-Lys62-Glu64 of the proximal ubiquitin. This unique surface area is found only in lys63 linked chains. Hence lysine residue of the proximal ubiquitin molecule involved in the linkage plays a crucial role in determining the specificity of these DUBs [27].

1.2.6 MINDY family

Motif interacting with ubiquitin (MIU) is a small UBD that can bind to monoubiquitin. During the investigation of this domain across different proteins, FAM63A was identified to contain MIU domain with selectivity for binding lys48 linked polyubiquitin

chain. Also, FAM63A contains DUF544 domain with a cysteine protease active site. The domain architecture of these proteases is distinct from known DUB family members and therefore named as MINDY (MIU-containing novel DUB family). MINDY family DUBs show specificity towards lys48 linked polyUb chains. Orthologs of FAM63 are also present in plants, budding yeast and Dictyostelium and notably yeast ortholog YPL191C (MIY1, MINDY in yeast) also retains the specificity to act on lys48 linked polyUb chains. Structural studies reveal a substrate-induced conformational remodeling in the active site, and the S1 site appears to bind ubiquitin in two alternative conformations. The biological function and regulation of this newly discovered DUB family is a new topic of future research [20].

1.3 UBL-specific proteases (ULPs)

These proteases act on ubiquitin-like modifiers such as SUMO and NEDD8. They are termed as ULP or SENP proteases and fall into a distinct class of proteases which do not resemble DUBs. The DeSI (deSUMOylating isopeptidase) family belong to PPPDE (permuted papain fold peptidases of double-stranded RNA viruses and eukaryotes) class of proteases [40], and they show little sequence similarity to ULPs or SENP protease class. Altogether they are referred to as UBL-specific proteases (ULPs) [21].

1.4 DUBs specificity

As mentioned earlier about the properties of ubiquitin and UBLs, these modifiers differ in their amino acid sequences but have a β-grasp fold. The mechanism by which these DUBs can recognize their substrate (ubiquitin/UBL) determines their specificity. In addition to the C-terminal region near di-glycine motif of ubiquitin, the surfaces on ubiquitin such as Ile44 patch (comprising Ile44, Leu8, Val70, and His68) and Ile36 (Ile36, Leu71, and Leu73) patch are known to mediate the DUBs interaction [4]. Except for Nedd8, the UBLs like SUMO, ISG15 and FAT10 exhibit minimal similarity to Ile44 patch. In fact from structural studies of protease SdeA (bacterial protease) with ubiquitin suicide substrate and USP CYLD with Met1 or Lys63-linked diubiquitin substrates have revealed that these proteases have no direct contacts with an Ile44 patch of ubiquitin [41,42]. Ubiquitin harbours additional protein binding elements like Phe4 patch (Gln2, Phe4, and Thr14), the TEK box (Lys6, Lys11, Thr12, Thr14, and Glu34), and the Asp58 patch (Arg54, Thr55, Ser57, and Asp58) [4]. Perhaps the most significant property of Ubiquitin and UBLs is their C-terminal flexible tail which is stabilized by the active site cleft of enzymes. The deconjugating enzymes heavily depend on the properties of C-terminal tail residues [43].

All DUBs contain a primary ubiquitin binding site called S1 (Fig.1.4 A) [19]. The interface between the S1 site and ubiquitin covers 20-40% surface of bound ubiquitin molecule. The catalytic domain engulfs the extended C-terminal tail of ubiquitin. These enzymes may also have S1′, S2, S2′ sites present in either cis or trans, which can bind to proximal ubiquitin moiety or the protein substrate (Fig.1.4 A and B). The enzymes which have only the S1 binding site can indiscriminately remove ubiquitin from their target protein. The property of a single ubiquitin binding site explains the nonspecificity of DUBs towards their substrates and the promiscuity of USP domain proteases which have a conserved S1 site [19].

The SAGA complex responsible for histone H2B deubiquitination for transcriptional regulation harbors Usp8 enzyme for deubiquitylation. During catalysis, the SAGA component Sgf11 establishes contacts with H2A/H2B proteins of nucleosomes. Here, the USP catalytic domain makes additional contacts with H2B and ubiquitin molecule indicating that Usp8 has evolved with histone (substrate) specific S1′ binding site [44].

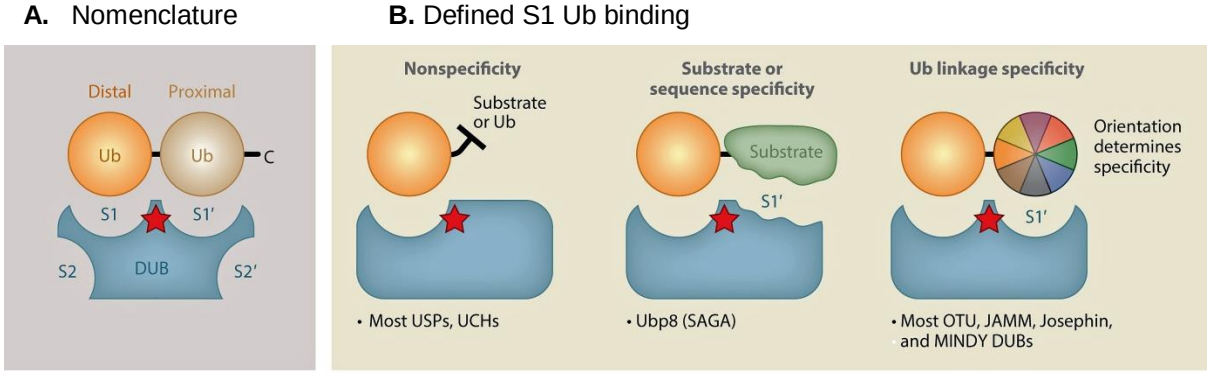


Figure 1.4 - Principles of substrate recognition and linkage specificity.
A) Basic nomenclature for DUB ubiquitin (Ub)-binding sites and a model diubiquitin substrate. The active site is indicated by a red star.
B) Defined S1 sites can be sufficient to recruit ubiquitinated target proteins or Ub chains. DUBs without an S1′ site are nonspecific, which applies to most USP family members. S1′ sites interacting with substrate protein features are the basis of substrate specificity. Ub linkage specificity is achieved by defined S1′ Ub-binding sites that bind the proximal Ub such that only one linkage point (indicated by different colors) can access the active site. Many members of OTU, JAMM, Josephin, and MINDY families are linkage specific. The figure is adapted from [19].

1.5 Mechanisms of DUBs regulation

Since DUBs are the key enzymes that regulate all ubiquitin-dependent processes, their abundance, localization, and catalytic activity are tightly controlled to ensure appropriate responses [45]. Mutations in DUBs or misregulation of their associated pathways have been linked to cancer, inflammatory diseases and neurodegeneration [46–48]. Mechanisms of DUBs regulation can be classified into two types; first is regulation by abundance and localization and second is regulation of DUB catalytic activity. The processes like post-translational modifications (PTMs), posttranslational processing, or accessory domains and binding proteins can act together for DUB regulation and thus interconnected with each other.

An example of DUB regulation by abundance is an A20 enzyme. Its levels in uninduced cells are low which rise substantially upon NF-κB activation. After induction of A20, its proteolytic activity can be regulated by paracaspase MALT1, which cleaves A20 between the N-terminal catalytic OTU domain and the C-terminal UBDs, thereby impairing A20 function [49]. The protease USP1 deubiquitinates the mono-ubiquitinated PCNA and help in preventing the activation of the error-prone translesion synthesis (TLS) of DNA. Upon exposure to ultraviolet (UV) irradiation, USP1 undergoes autoprocessing after a Gly-Gly motif present within the USP domain resulting in its degradation. Whether the residues adjacent to the Gly–Gly motif or putative UBL domain upstream of diglycine motif contribute to the auto-cleavage consensus site has not been determined [50].

A systematic analysis of DUBs localization in fission yeast or mammalian cells has revealed that DUBs vary in their distribution in the cell and are localized specifically in cell compartments to perform their function [51,52]. DUB localization is achieved by targeting domains to interact with distinct cellular membranes, organelles, cytoskeletal components, or through localization signals, and protein interaction domains that recruit DUBs to defined complexes [19]. Also, PTMs (Post Translational Modifications) are found to regulate DUB localization for example; phosphorylation of OTUB1 and ATXN3 by casein kinase 2 and phosphorylation of USP10 by ATM kinase (ataxia-telangiectasia mutated), results in their nuclear localization [19].

USP8 enzyme activity is inhibited due to its phosphorylation and association with 14-3-3 protein, whereas dephosphorylation of USP8 in M phase enhances its activity [53]. The DUB VCIP135 is essential for p97/p47-mediated post-mitotic Golgi membrane fusion. In early mitosis, phosphorylation of VCIP135 by Cdk1 at a single residue (S130), inactivates the enzyme activity and thereby regulate the Golgi assembly [54].

DUBs can also be ubiquitinated when complexed with E3 ligases, and such covalently attached ubiquitin can compete for binding to ubiquitinated substrates. The DUBs like UCH-L1, USP30, ATXN3 are reported to get ubiquitinated and may regulate their catalytic activity [19]. USP25 is upregulated during viral infections and involved in interleukin signaling. It has three N-terminal UBDs which modulate DUB activity, and the enzyme is activated by monoubiquitination of Lys99 located at the beginning of a UIM. The same lysine residue can be sumoylated. It is postulated that alternative conjugation of ubiquitin (activating) or SUMO (inhibitory) to Lys99, may promote the interaction with distinct intramolecular regulatory domains [55]. A reactive cysteine residue present at the active site of cysteine proteases is susceptible to oxidation. A majority of DUBs belong to this class of proteases and oxidation by reactive oxygen species (ROS) has been shown to inactivate members of the OTU, USP, and UCH family enzymes *in vitro* and *in vivo* [56–58]. The DUBs like USP5 and USP7/HAUSP are known to be allosterically activated. The ubiquitin-associated (UBA) and ZnF-UBP domains present in USP5 or five UBL domains present in USP7 are known to enhance catalytic activity by allosteric modification [19].

1.6 Ubiquitin-like proteins (UBLs)

There is a class of proteins which have a characteristic ubiquitin-like fold (ββαββαβ, i.e., β grasp fold) a five-stranded β -sheet with a single helix on top, common to ubiquitin and ubiquitin-like proteins (Fig.1.6). Most of these UBLs utilize the E1/E2/E3-like conjugation machinery and require proteolytic processing for their activation. Selective UBLs also get conjugated to target proteins or lipids and perform crucial roles besides protein degradation. Among them are, Nedd8 (Neural precursor cell expressed, developmentally downregulated 8), ISG15 (Interferon-stimulated gene 15), SUMO (Small ubiquitin-related modifier), FAT10 (HLA-F adjacent transcript 10), Atg8 and Atg12 (Ubiquitin related autophagy protein), URM1 (Ubiquitin-related modifier 1), UFM1 (Ubiquitin fold modifier 1), and Hub1 (Homologous to ubiquitin). The cellular function of these UBLs is discussed below. The summary of roles of mammalian UBLs in cellular processes and their homology with ubiquitin is listed in table 1.6.

Table 1.6- List of mammalian ubiquitin-like proteins [59]

Ub/UBL	*S. pombe*	Identity (%)	C-terminal processing	Protease	Function
Ubiquitin	Ubiquitin	100	Yes	~80 DUBs	Multiple functions dependent on linkage
Nedd8	Ned8	58	Yes	CSN5, NEDP1	Regulate Ub conjugation via CRLs, cell cycle
ISG15	–	28/37	Yes	UBP43	Antiviral immunity, IFN Inducible
*SUMO	SUMO	~13	Yes	SENPs	Alter interactions, localization, conformation
FAT10	–	27/36	No	–	Ub-dependent proteasomal degradation, immunoregulatory role
Ufm1	–	23	Yes	UfSP1, UfSP2	Erythroid and megakaryocyte development
Urm1	Urm1	17	No	–	tRNA thiolation and oxidation-induced protein modification
Atg12	Atg12	12	No	–	Autophagy, mitochondrial homeostasis
*Atg8	Atg8	~9 to ~14	Yes	Atg4	Autophagy
UBL5	Hub1	22	No	–	Noncanonical splicing, unlikely to be conjugated

(*) - There are isoforms of SUMO and Atg8 present in humans; therefore, % identity is shown as average or a range. For ISG15 and FAT10 the identities for each of the two Ub-related domains are listed. (−) Blank spaces indicates the absence of an ortholog in *S. pombe* or unidentified enzymes

1.6.1 Nedd8

The Nedd8 coding gene was first identified as one of the ten Nedd (Neural precursor cell expressed, developmentally downregulated) genes abundantly expressed in embryonic mouse brain [60]. Among UBLs, Nedd8 protein has the closest sequence identity with ubiquitin (58%). It is also synthesized as a precursor and processed at the C-terminal di-glycine motif by NEDP1 (DEN1, SENP8) or UCHL3 (the protease also acts on ubiquitin). Nedd8 gets conjugated to protein substrate by the sequential action of a Nedd8-activating enzyme (a heterodimer composed of NAE1 and Uba3), the E2 enzymes Ubc12 (Ube2M) or Ube2F, and a few E3 ligases [59]. Nedd8 forms conjugates with members of the Cullin family of proteins which are essential for assembly of Cullin-RING E3 ligases (CRLs). These

ligases function in ubiquitination and proteasome degradation pathways [60]. The CRLs activity is involved in the regulation of cytokinesis, oxidative stress response pathways, DNA replication, adenoviral replication, Hedgehog signaling pathways and degradation of oncoproteins [61].

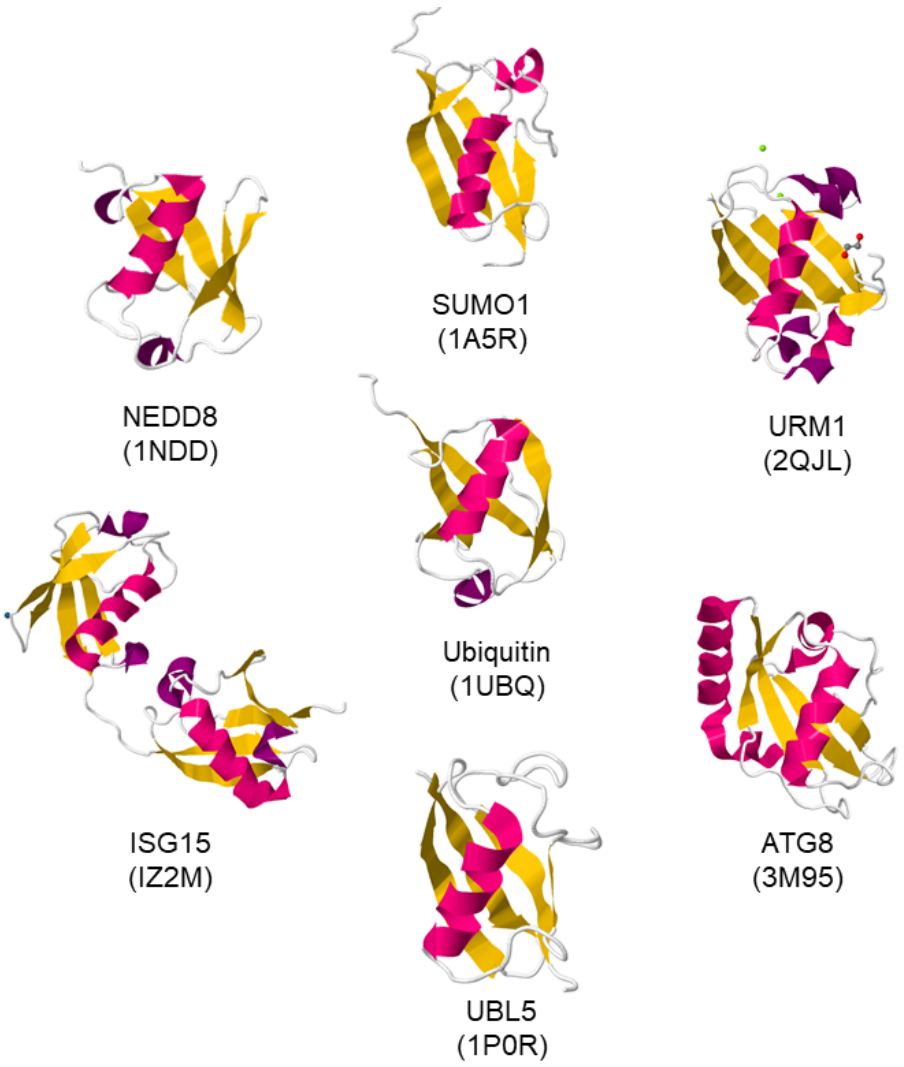


Figure 1.6 – The β-grasp fold shared by Ubiquitin and UBLs

Structural comparison of Ubiquitin, NEDD8, SUMO1, URM1, ATG8, UBL5 and ISG15. The respective PDB codes are mentioned below each UBL structure.

The transcription factor p53 also undergo neddylation by E3 ligase MDM2 which inhibits p53 transcriptional activity. [62]. Neddylation is mainly reversed by DEN1/NEDP1 enzymes [62]. Other de-neddylating enzymes are CSN5, a subunit of COP9 signalosome, UCHL1, UCHL3, Otubain-1, USP21, PfUCH54, and Ataxin3. Some pathogens have evolved with

proteases which can act on both Nedd8 and Ubiquitin, for example, *Chlamydia trachomatis*, Epstein-Barr virus, *Plasmodium falciparum* [59,61].

1.6.2 ISG15

The ISG15 coding gene was identified among hundreds of interferon-stimulated genes (ISGs) induced by Type I interferon which modulates defense mechanism of the host during pathogenic infections. ISG15 have two ubiquitin-like domains with ~30% amino acid sequence homology to ubiquitin, linked together by a hinge [63]. In addition to the Type I interferon stimulation, ISG15 is induced by viral and bacterial infections, lipopolysaccharides (LSPs), retinoic acid and some genotoxic stressors [63]. This gene is absent in yeast, *C. elegans*, and *D. melanogaster* and found in higher eukaryotes having IFN signaling [64]. Conjugation of ISG15 is termed as ISGylation which involves sequential enzymatic cascade that carries out ISG15 conjugation of the substrate [59]. The process of ISGylation can be reversed by ubiquitin carboxy-terminal hydrolase 18 (USP18), UBP43 and a viral Ovarian Tumour Domain (OTU) containing proteases which target both ubiquitin and ISG15. ISGylation of viral proteins can reduce viral replication by disrupting viral-host protein interactions, or it can inhibit oligomerization of viral proteins essential for viral replication. For example, ISGylation of non-structural protein-1 (NS1/A) of Influenza A virus (IAV) at distinct sites disrupts its interaction with host proteins limiting its ability to modulate host antiviral response [63]. ISGylation can directly affect viral protein function, such as ISGylation of human cytomegalovirus (HCMV) pUL26 protein alters its stability thereby inhibiting TNFα-induced NF-κB activation/signaling [65,66]. The ISGylation of viral protease 2APro hampers its activity on translation initiation factor eIF4G1 which in turn reduces viral replication [67]. Overall ISG15 is a major player in mediating antiviral response which modulates both viral and host proteins, regulates cytokines release and prevents pathogen invasion [68,69].

1.6.3 SUMO

The term Small Ubiquitin-like Modifier (SUMO) was given after a protein called GMP1 (GAP modifying protein 1) which was discovered to be covalently attached to GTPase activating protein RanGAP1 [70,71]. The members of the SUMO family are present in all eukaryotic kingdoms and highly conserved from yeast to humans [72]. *S. cerevisiae* contains a single *SMT3* gene essential for viability and the temperature-sensitive mutants of the SUMO conjugation pathway show cell cycle defects. In *S. pombe*, Δ*pmt3* (SUMO) or Δ*hus5*

(Ubc9, a SUMO-conjugating enzyme) strains show hypersensitivity to DNA damage and chromosome mis-segregation [72].

In vertebrates there are 4 SUMO coding genes designated as SUMO1−4. The isoforms SUMO2 and SUMO3 are 97% identical therefore collectively referred to as SUMO2/3, and they are ~50% identical to SUMO1. The isoform SUMO4 is ~87% similar to SUMO2, but its role in sumoylation pathway is unclear [59]. A prominent difference between the SUMO family and other UBLs is the presence of a flexible N-terminal extension in SUMO. Although the primary sequence identity between SUMO1 and ubiquitin is around 18%, the nuclear magnetic resonance (NMR) structure shows a similar β grasp fold shared between the two proteins [73]. A heterodimeric E1 enzyme AOS1-UBA2 activates SUMO by forming a thioester bond with active site cysteine residue present on UBA2 subunit. A SUMO-specific E2-conjugating enzyme UBC9 physically interacts with almost all known SUMO substrates [72]. The E3 ligases for SUMO are not yet identified.

SUMO is synthesized as a precursor and requires processing at the C-terminus by ULPs (Ubiquitin-like protein processing enzymes) or SUMO-specific proteases for activation. The two enzymes Ulp1 and Ulp2 from yeast can catalyze C-terminal processing of SUMO and can break isopeptide bond between SUMO and target protein (a process known as desumoylation). In humans, at least 7 ULPs were identified based on the homology with the ULP domain of yeast ULPs and termed as SENPs or SUSPs for sentrin/SUMO-specific proteases [74]. As the mammalian cells have three SUMO isoforms (SUMO-1, -2 and -3), each SENP enzyme could be associated with specific function and localization.

Sumoylation of several transcription regulators like HIPK2, TEL, c-jun, p53 is reported to alter their transcriptional potential. Whether this downstream effect occurs due to altered localization or structure requires more studies [75–77]. The SUMO-1-RanGAP1 conjugate can interact with Ran-binding protein-2 (RanBP2), a protein associated with 'Nuclear pore complex' (NPC). SUMO induces structural changes in RanGAP1 that allows its binding with RanBP2. Sumoylation of promyelocytic leukemia protein (PML) induces relocalization of a transcriptional corepressor Daxx into nuclear bodies causing its inactivation [72]. Contrary to this, the PML sumoylation directs p53 to nuclear bodies and leads to stimulation of its transcriptional and pro-apoptotic activity [78].

In NF-κB pathway, the transcription factor NF-κB is kept inactive in the cytosol by binding of the inhibitor IκBα. Upon stimulation by effector such as tumor necrosis factor (TNF), IκBα gets phosphorylated, ubiquitinated and degraded by the proteasome. Here,

SUMO conjugation to IκBα at the same lysine residue used by ubiquitin inhibits its degradation and thus NF-κB function [79]. The sumoylation of ubiquitin E3 ligase Mdm2 also prevents its self-ubiquitination and degradation via proteasome [80].

1.6.4 FAT10

FAT10 is a cytokine-inducible modifier termed as HLA-F adjacent transcript 10 encoded at the MHC gene locus. The gene is induced by pro-inflammatory cytokines IFN-γ and TNF-α and reported to function in immune defense, proteasomal degradation, and cancer development [81–83]. Similar to ISG15, it also has two tandem ubiquitin-like domains and a di-glycine motif at the C-terminus for conjugation with the substrates. Activation and conjugation of FAT10 to specific substrates is mediated by an E1-type enzyme Uba6 (UBE1L2, E1-L2, or MOP-4) and USE1 (Uba6-specific E2 enzyme) where both the enzymes show dual specificity for ubiquitin and FAT10. The activated FAT10 is transferred to USE1 and becomes active for FAT10ylation of substrates. USE1 can undergo auto-FAT10ylation in cis, mainly on Lys323. When FAT10 conjugated substrate reaches proteasome, a portion of the auto-FAT10ylated USE1 also gets degraded. Therefore, this mechanism limits the FAT10 conjugation pathway by a negative feedback loop system [59,83].

FAT10 non-covalently interacts with NEDD8 ultimate buster-1 long (NUB1L), where both the proteins can independently associate with the 26S proteasome. NUB1L is proposed to make a conformational change in the proteasome which facilitates binding and degradation of the FAT10 conjugated substrates [81]. The role of FAT10 in MHC antigen presentation was demonstrated in HeLa cells [82]. The FAT10 decorated cytosolic *Salmonella typhimurium* are targeted for autophagy similar to ubiquitin, suggesting its role in antimicrobial defense [84]. Several publications support the notion that FAT10 is a proto-oncogene and it is found upregulated during inflammatory processes involved in the development of cancer [83]. Further, the FAT10 expression mediates a positive feedback loop which enhances NF-κB activation and may contribute to the malignant properties of FAT10. However, the FAT10 downstream targets and consequences of their interactions leading to NF-κB activation are not known [83].

1.6.5 ATG8 and ATG12

ATG12 and ATG8 are ubiquitin-like proteins which were identified among 30 autophagy-related genes (Atg). Autophagy is induced during low nutrient availability, stress

responses, development, immunity, and inflammation. [59,85]. Each component of the Atg12 system is found in mice and humans, with functional conservation to that of yeast [85]. Atg12 have C-terminal glycine residue and do not require precursor processing for activation and conjugation with its only target Atg5. The E1-like enzyme Atg7 activates Atg12 via formation of thioester bond and then transferred to an E2 enzyme Atg10. Finally, it gets conjugated to the target protein Atg5 at Lys149 through an isopeptide bond. An E3 enzyme for Atg12–Atg5 conjugation is not reported in the literature. The Atg5 protein interacts with Atg16 forming a multimeric complex of Atg12–Atg5-Atg16 through the homo-oligomerization of Atg16 [85].

The single *ATG8* gene present in *S. cerevisiae* has several paralogs present in multicellular animals. The *Sc*Atg8 protein has a single C-terminal Arg117 residue which is proteolytically removed by a cysteine protease Atg4 to expose Gly-116. The processed Atg8 is activated by E1-like enzyme Atg7 and further transferred to the E2-like enzyme Atg3. In the final step of lipidation, Atg8 is conjugated via its Gly-116 to phosphatidylethanolamine (PE) through an amide bond and this Atg8–PE conjugate is found associated with membranes. Unlike Atg12–Atg5 conjugation, Atg8–PE conjugation can be reversed by Atg4 mediated cleavage to release free Atg8 [85].

1.6.6 URM1

The ubiquitin-related modifier 1 (URM1) is a 99 amino acid protein with a β-grasp fold & have a typical di-glycine motif which operates in ubiquitin-like conjugation pathways termed as URMylation and tRNA thiolation [86]. Urm1 is conserved from *S. cerevisiae* to *Homo* sapiens [59], and it was first identified from the yeast database search for proteins similar to prokaryotic sulfur carrier proteins MoaD and ThiS found in prokaryotic molybdopterin (MPT) and thiamine (Vitamin B1) biosynthesis pathways respectively. This finding provided evidence for an ancestral link between the prokaryotic origin of eukaryotic ubiquitin/UBL modification pathways [87,88]. The MPT and thiamine are synthesized by transfer of sulfur from donor MoaD & ThiS to MPT precursor Z and thiazole respectively. Both the sulfur donors MoaD and ThiS, contain a C-terminal diglycine motif which is derivatized by MoeB or ThiF enzymes to introduce sulfur atom via thiocarboxylate linkage [89].

Urm1 is activated by forming an acyl-adenylate between the C-terminus of Urm1 and AMP and further derivatized by the E1-like enzyme Uba4 (molybdenum cofactor synthesis 3, MOCS3 in *H. sapiens*) by forming a thiocarboxylate linkage [59]. Thiocarboxylated URM1

interacts with the thiouridylase Ncs6 or ATPBD3 (ATP binding domain 3) in humans, the enzymes involved in tRNA thiolation. Such modifications, are although not essential for the viability of cells but confer stability to the modified tRNA and help in translational fidelity and efficiency [59]. Among the known targets of URM1 are, the peroxiredoxin Ahp1 gets URMylated and is reported to affect its function in ROS protection. URM1 may modify the yeast Uba4 itself, and such auto URMylation may have feedback loop regulation but requires further studies [86].

1.6.7 UFM1

The Ubiquitin-fold modifier-1 is a 9.1 kDa protein identified as a protein modifier with only 16% sequence identity to ubiquitin. However, it shows a similar tertiary structure and enzymatic cascade like ubiquitin. Ufm1 is synthesized as a precursor (proUfm1) which is cleaved at the C-terminus to remove Ser-Cys dipeptide. The processing exposes Gly83 residue necessary for conjugation to a target substrate [90]. In human cells, the proUfm1 processing and deconjugation of UFMylated proteins are carried out by two Ufm1-specific proteases (UfSPs) [91]. The Ufm1 conjugation to target proteins is mediated by an enzymatic cascade conserved in eukaryotes except for yeast.

Some of the reported target substrates of Ufm1 are Ufm1-binding protein 1 (UfBP1), activating signal cointegrator-1 (ASC1) [92]. UfBP1 is predominantly localized in the ER and is known to mediate protein-protein interaction and formation of several multiprotein complexes via its PCI domain: for example COP9 signalosome (CSN), 26S proteasome, and the eukaryotic translation initiation factor 3 (eIF3) complexes [93]. ASC1 is a transcriptional coactivator of estrogen receptor-a (ERa) and also a tumor suppressor known to activate p53, induce apoptosis, and suppress NF-κB signaling [94]. Ufm1 modifications are associated with endoplasmic reticulum (ER) stress, hematopoiesis, fatty acid metabolism, and G-protein coupled receptor (GPCR) biogenesis. The defects in Ufm1 cascade are implicated in human diseases like cancer, ischemic heart diseases, and schizophrenia. Importantly, Ufm1 is essential for embryonic development, erythroid differentiation, and hematopoietic stem cells survival [92,94,95].

1.6.8 HUB1

Hub1 (homologous to ubiquitin-1) also termed as UBL5 or beacon in mammals. It is an unconventional ubiquitin-like protein because, i) It lacks a typical di-glycine (GG) motif

found in archetypal UBLs. Instead, it possesses a C-terminal double tyrosine (YY) motif followed by a non-conserved amino acid residue, ii) The enzymatic cascade required for activation and conjugation of Hub1 is absent or not reported yet in any organism. Hub1 is 73 amino acid residue protein (Mol. Wt. 8.5 kDa) which is highly conserved in eukaryotes. In *S. pombe*, the conditional mutants of Hub1 showed moderate splicing defects and were found to non-covalently associate with spliceosomal U4/U6.U5 tri-snRNP protein Snu66 [96,97]. Hub1 regulates the alternative splicing of *S. cerevisiae* gene *SRC1* by binding non-covalently (via surface-I) to Snu66 [98]. The Hub1-Snu66 interaction depends on an 18–19-amino acid long helical Hub1 interaction domain (HIND) found in Snu66 of yeast, vertebrates, and also in plant snRNP protein Prp38 [98]. Recently, Hub1 is shown to bind via its surface-II with DExD/H-box helicase Prp5. This interaction helps in relaxing splicing fidelity, necessary for splicing of suboptimal introns in *S. cerevisiae*. Hub1 also stimulates the ATPase activity of Prp5 thereby improving the splicing efficiency [99]. It appears that Hub1 operates through multiple surfaces to regulate pre-mRNA splicing. In *C. elegans* UBL5 is found to regulate the mitochondrial unfolded protein response by inducing chaperone genes expression [100]. In human cells, Hub1/UBL5 is essential for viability, and prolonged depletion of Hub1 showed splicing speckle abnormalities, partial nuclear retention of mRNAs, and mitotic catastrophe [101]. UBL5 maintains the integrity of the Fanconi anemia (FA) pathway by increasing the stability of FANC1 protein [102].

1.7 RNA splicing

Most of the genes in eukaryotes are transcribed as pre-mRNAs which contain non-coding sequences (introns) and coding sequences (exons). Introns are removed during the process of pre-mRNA splicing, and exons are joined to produce mature mRNAs (Fig.1.7). Human genes contain multiple introns where splicing plays a crucial role in gene expression. [103]. Pre-mRNA splicing reactions are catalyzed by a dynamic ribonucleoprotein (RNP) machine known as spliceosome. It consists of five different RNP subunits namely U1, U2, U4/U6, and U5 together with several associated protein cofactors. To distinguish them from other cell RNPs like ribosomes they are termed as small nuclear RNPs [104]. Introns are removed by two consecutive transesterification reactions catalyzed by the spliceosome. The nucleotides (nts) at the intron termini and their flanking exon regions possess splicing signals which are recognized by spliceosome for accurate and efficient splicing. A donor splice signal at the 5′ of the intron (5′ ss), branch point (BS) adenosine in an intron, polypyrimidine

tract preceding 3′ end of intron and acceptor splice signal at 3′ end of intron (3′ ss) defines the characteristics of an intron.

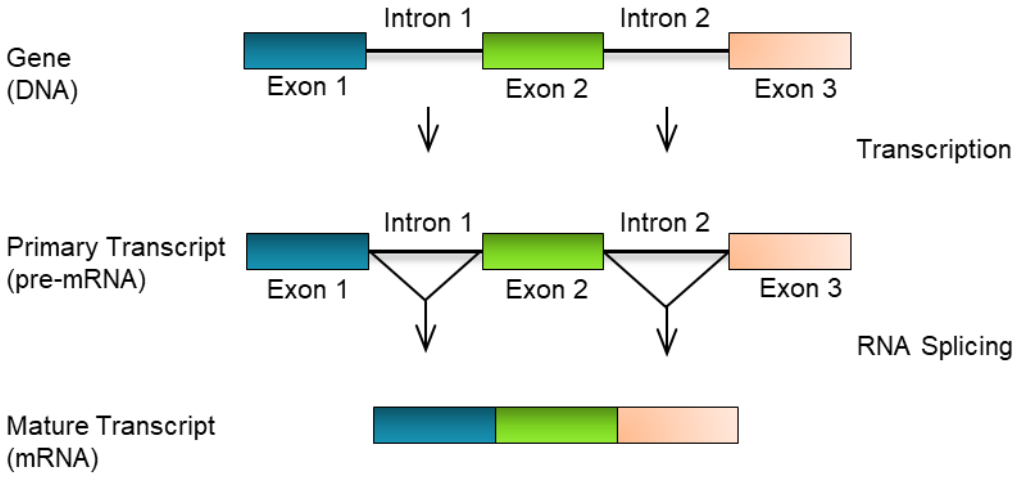


Figure 1.7 - RNA splicing: Introns from pre-mRNAs are removed by the process of splicing, followed by ligation of exons to generate translatable mRNAs. This process is facilitated by an RNA-protein complex called spliceosome.

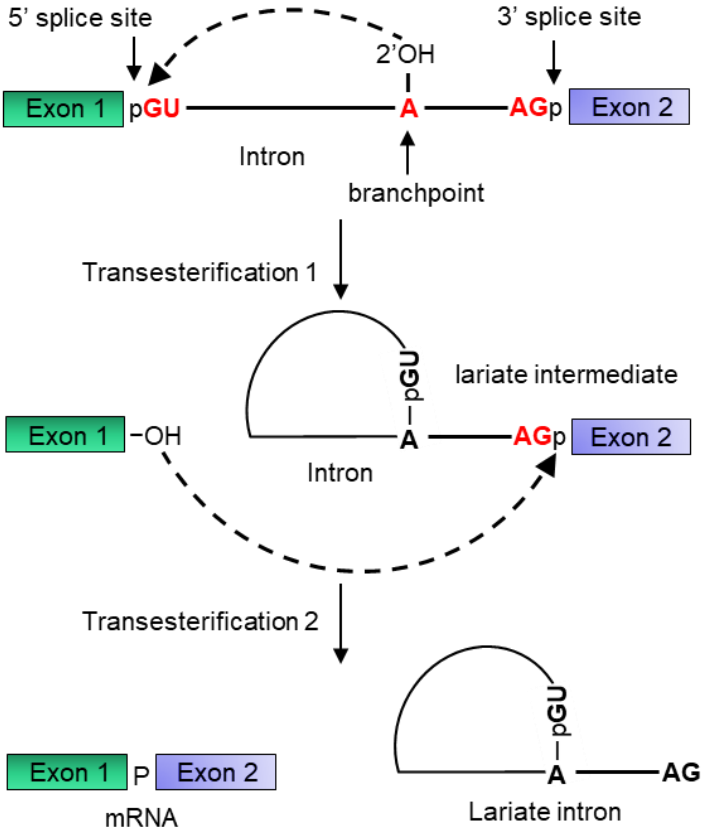


Figure 1.7.1 - Systematic steps of transesterification reactions during splicing: During pre-mRNA splicing, the two transesterification reactions at step I and step II result in the production of the spliced product mRNA and intron-lariat.

In step-I, the 2′ OH group of the branch site adenosine carries out a nucleophilic attack on the 5′ ss resulting in the cleavage between 3′ end of an exon and 5′ end of the intron. This results in the formation of 2′–5′ phosphodiester bond between the branch adenosine and 5′ end of the intron leading to the formation of an intron-lariat structure (Fig.1.7.1).

In step-II, the 3′ ss is attacked by the 3′ OH group of the 5′ exon followed by 5′ and 3′ exons ligation. The outcome of these steps is the formation of mRNA and release of the intron (Fig.1.7.1) [104]. The mRNA undergoes several processing steps like 5′ capping and the addition of a polyA tail before it is transported to the cytoplasm for translation [105].

1.7.1 Assembly of spliceosome

The main components of the spliceosome are U1, U2, U4/U6, and U5 snRNPs, each one consisting of a snRNA and a common set of seven Sm proteins (B/B′, D3, D2, D1, E, F, and G) and a variable number of particle-specific proteins. The U2 dependent spliceosome is responsible for removing the majority of U2-type introns. A subset of eukaryotes contains less abundant U12 dependent spliceosome made up of U11, U12, U4atac, U6atac, and U5 snRNPs, involved in the removal of U12-type introns [106].

In events where intron length does not exceed ~200-250 nucleotides the spliceosome initially assembles across intron [107]. The 5′ end of the U1 snRNA is complementary to the 5′ ss of the intron. Assembly of spliceosome begins with the ATP-independent binding of the U1 snRNP through base-pairing interactions usually considered to take place between the positions −3 to +6 of the exon/intron junctions with consensus (CAG/GURAGU) sequence. In higher eukaryotes, these interactions are stabilized by serine-arginine-rich (SR) proteins and proteins of the U1 snRNP.

In addition to this, binding of SF1/BBP protein and U2 auxiliary factor (U2AF 65-kDa subunit) to the branch point and the polypyrimidine tract respectively marks the early phase of spliceosome assembly called as E complex formation [108,109]. Another 35 kDa subunit of U2AF contacts the 3′ ss and plays a role in 3′ ss recognition [110]. In a subsequent step, the U2 snRNP stably associates with the branch site where U2-associated SF3a and SF3b proteins play an essential role in binding [111].

The U4/U6.U5 tri-snRNP, pre-assembled from U5 and U4/U6 snRNPs is then recruited to generate the pre-catalytic B complex. The pre-assembled U4/U6 base-pairing interaction is unwound, allowing the U6 snRNA to form new associations with the U2

snRNA and the 5′ splice site, provoking the destabilization of U1 and U4. This causes rearrangements in RNA-RNA and RNA-protein interactions leading to release of the U1 and U4 snRNPs.

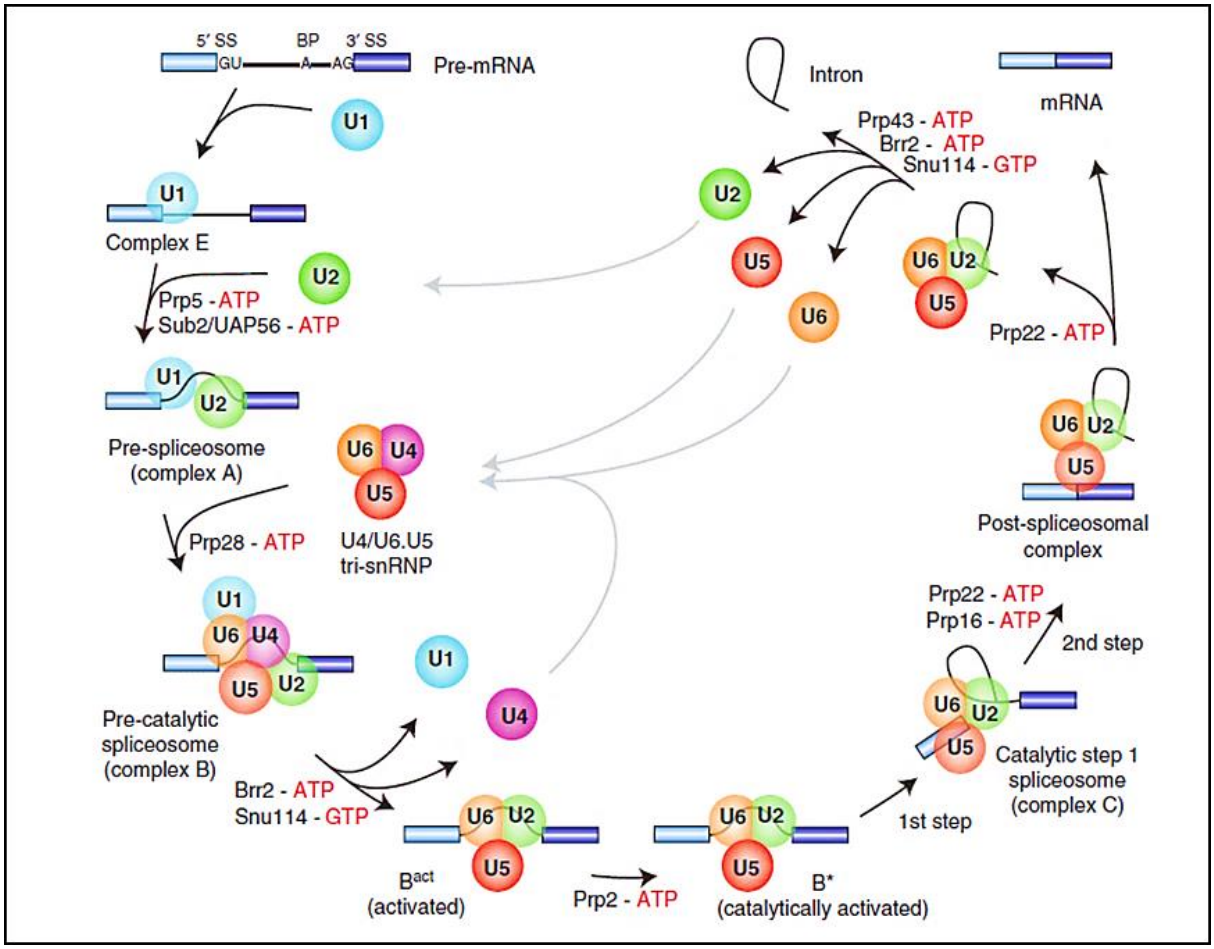


Figure 1.7.2 – Spliceosomal assembly and disassembly pathway of U2-dependent spliceosome. The ordered interactions of the snRNPs (indicated by circles), but not those of non-snRNP proteins, are shown. The various spliceosomal complexes are named according to the metazoan nomenclature. Exon and intron sequences are indicated by boxes and lines, respectively. The stages at which the evolutionarily conserved DExH/D-box RNA ATPases/helicases, Prp5, Sub2/UAP56, Prp28, Brr2, Prp2, Prp16, Prp22, and Prp43, or the GTPase Snu114, act to facilitate conformational changes are indicated. The figure is adapted from [104]

Further addition of the Prp19-associated NineTeen Complex (NTC) to the spliceosome gives rise to catalytically activated complex B (complex B^{act}/ B*). There are multiple RNA helicases (Brr2, 114 kDa U5 small nuclear ribonucleoprotein component (Snu114) and Prp2) which are required for the activation of complex B [112]. The complex B* catalyzes the first transesterification reaction of the splicing which yields a C complex containing the free exon-1 and the intron-exon-2 lariat intermediate. The C complex

undergoes additional ATP dependent rearrangement and catalyzes the second transesterification reaction of splicing which is dependent on Prp8, Prp16 and synthetic lethal with U5 snRNA 7 (Slu7). The second step results in a post-spliceosomal complex that contains the lariat intron and spliced exons. It also releases the U2, U5, and U6 snRNPs to be recycled for additional rounds of splicing [105]. The schematic of the spliceosome assembly and dissociation is shown in Fig.1.7.2 [104].

The non-snRNP protein complex called Prp19 complex (Prp19C) or NineTeen Complex (NTC) plays a crucial role in splicing reaction by regulating the formation and progression of essential spliceosome conformations. The complex is named after its founding member Prp19 which was identified in a screen for DNA damage sensitive mutants [113]. The NTC is made up of eight core proteins together with 19 associated proteins in *S. cerevisiae* and more than 30 proteins in higher eukaryotes [114]. The NTC promotes new interactions between U5, U6 snRNAs and the pre-mRNA. It also destabilizes interactions between the U6 snRNA and Sm-like (Lsm) proteins during complex C formation [114]. Prp19 contains an N-terminal U-box/RING finger domain and has an E3 ubiquitin ligase activity [115]. In humans, Prp19 targets Prp3 (a U4 snRNP subunit) by adding K63-linked ubiquitin chain which enhances its interaction with Prp8, a core protein of the tri-snRNP. This interaction is important for stabilization of the U4/U6.U5 tri-snRNP complex [116]. Ubiquitylated hPrp3 is recognized by hPrp24, a U4/U6 snRNP component, and deubiquitylated by hUsp4 before spliceosome activation [116].

1.7.2 Exon and intron definition in pre-mRNA splicing

One of the crucial tasks of the spliceosome is to identify the bonafide splice site among many pseudo sites present in any pre-mRNAs. The 5′ and 3′ ss consensus motifs are degenerate [117] Therefore, the relative strength or similarity of splice site to the consensus sequence determines the fate of snRNP binding [118]. However, the splice sites recognized commonly by splicing machinery are highly divergent from the optimal sequences, and yet they are efficiently utilized for splicing. In order to recognize the real sites of splicing, many cis-acting elements present in exons and their nearby intronic regions within approximately 50 nucleotides from the exon boundaries [119] contribute to the recognition by the spliceosome.

Vertebrate genes are characterized by short exons (~170 nts length) separated by considerably longer introns (~1900 nts length) [120,121]. When exons are short, and introns are long the splicing machinery is more likely to form across exon, the process is termed as

'exon definition' [122]. During exon definition, the U1 snRNP binds to the 5′ss downstream of an exon and promotes the association of U2AF with the polypyrimidine tract/3′ss of an intron located upstream of the exon. The U2 snRNP is then recruited to the BS, and further spliceosomal assembly takes place.

The exon defined complex is stabilized by recruitment of proteins of the SR protein family by splicing enhancer sequences within the exon (ESEs) [123,124]. In lower eukaryotes like yeast the genome architecture is characterized by small introns and large exons where splicing machinery is assembled across intron first, called an intron definition model. During intron definition, U1 snRNP complexed to the 5′ss interacts with U2 snRNP and associated factors are assembled on the branch point, and 3′ss across the short introns and splicing is initiated [125].

1.7.3 Alternative splicing

The exon can be included or excluded from the final mRNA leading to generation of different isoforms from a single gene. The differential inclusion of exons in the final processed RNA product by splicing of a pre-mRNA is called alternative splicing (AS). Alternative splicing is the basis for difference between the estimated number of 24,000 protein-coding genes in the human genome and the 100,000 different proteins postulated to be synthesized [126]. A high throughput study has revealed that >90% of human genes undergo AS [127]. It is mainly responsible for maintaining levels and tissue specificity of gene expression, if disrupted may lead to disease [128,129]. Studies have suggested that AS complexity differs among vertebrate lineages and it has diverged to become species-specific [130].

1.7.4 Mechanism of alternative splicing regulation

Some of the studied mechanisms of alternative splicing are exon-skipping, intron-retention (IR), non-constitutive 5′ ss and/or 3′ ss usage, alternative promoters, alternative polyadenylation sites, splicing silencers and enhancers which contribute to various AS events [131–134]. Intron retention is most common in lower metazoans, fungi, and protozoa, whereas, exon skipping is more prevalent in higher eukaryotes and it might contribute most to phenotypic complexity [135]. Alternative 5′ and 3′ splice site usage might represent an intermediate evolutionary stage in alternative splicing [136]. The schematic of mechanisms of AS are given in (Fig.1.7.4).

In AS, the splice site selection by extrinsic, non-spliceosomal RNA-binding proteins (RBPs) plays a significant role. These are of three types, i.e., classical/canonical hnRNPs (heterogeneous nuclear ribonucleoproteins), SR proteins (serine/arginine-rich), and tissue-specific RNA-binding proteins like Nova, neuronal PTB/hnRNPI, the Rbfox family, and the muscleblind/CELF family protein [131]. SR proteins can bind RNA sequence motifs present in either exon or intron and act as splicing enhancers or silencers respectively [137,138]. The hnRNP proteins also have RNA sequence-specific binding activity and often work as splicing silencers. However, the splicing factor hnRNPL can activate splicing [139].

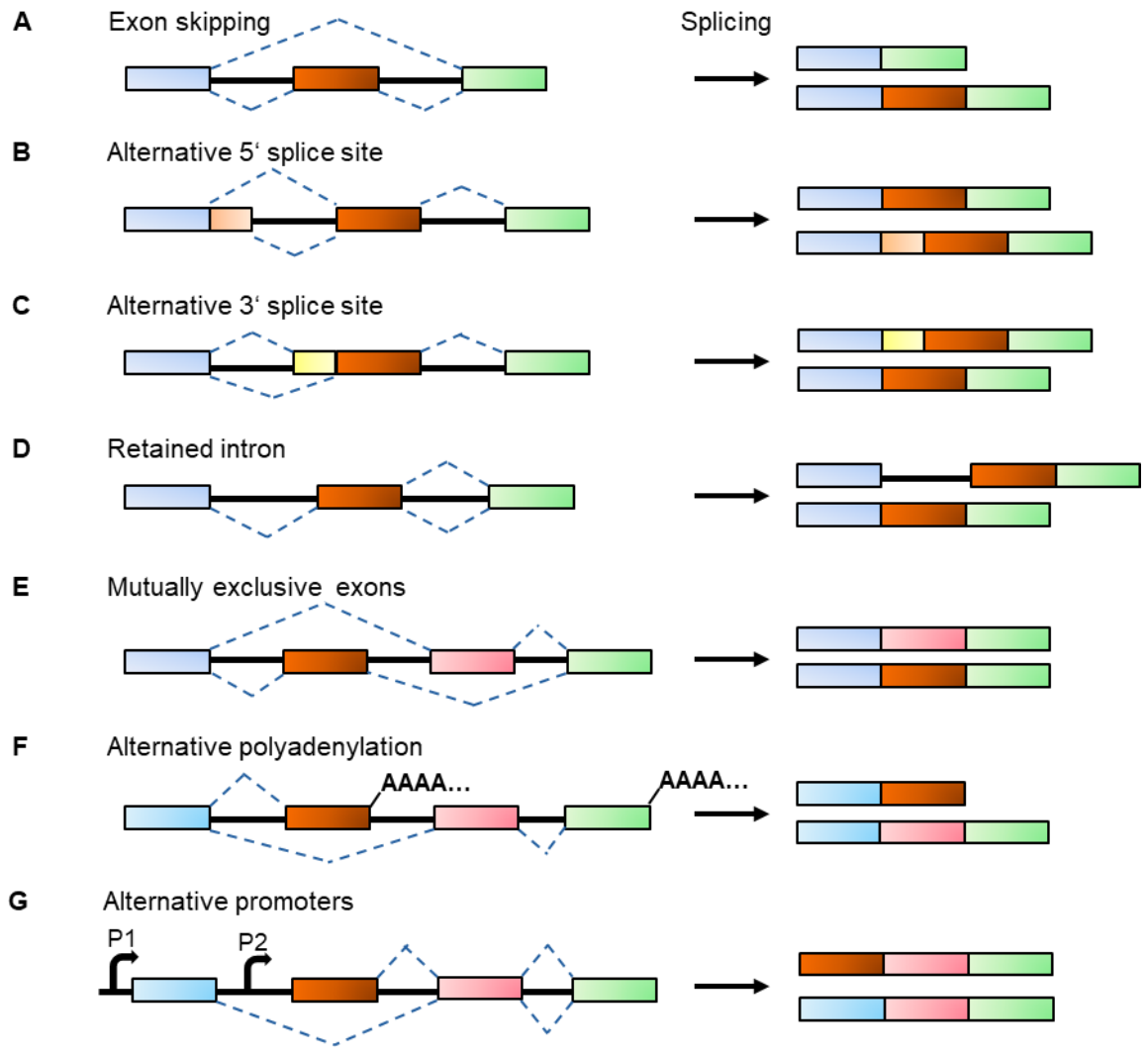


Figure 1.7.4 - Mechanisms of alternative splicing
Alternative splicing where one pre-mRNA gives rise to more than one mRNAs is mediated by multiple ways like skipping of exons, by using alternative 5' and 3' splice sites, intron-retention, the inclusion of mutually exclusive exons, alternative promoters and polyadenylations at different sites. Alternative 5' splice sites (in **B)** and alternative 3' splice sites (in **C**) are represented by cream and yellow coloured small rectangles respectively. P1 and P2 mark the presence of alternative promoters in transcripts.

Human hnRNPA1 splicing factor have different modes of action. a) its binding to exonic or intronic splicing silencer elements (ESS or ISS) represses the exon inclusion by steric action. b) the cooperative binding and "spreading" of hnRNPA1 proteins to adjacent lower-affinity sites through recruitment at a higher-affinity binding sites, and c) interaction of hnRNPA1 proteins bound to ISS (Intronic splicing silencers) elements present on both sides of an alternative exon, results in loop formation and exclusion of the exon [140–142]. The splicing factor hnRNPL is reported to regulates AS by interfering with 3′ splice-site recognition by $U2AF^{65}$ [143]. The splicing activator protein TIA-1 binds U-rich intronic regulatory elements adjacent to 5′ splice sites and interacts directly with the U1 snRNP C protein to promote 5′ splice-site usage [144].

RNA-RNA secondary structures may also control the alternative splicing [145]. Studies with chemical probes to monitor RNA secondary structure coupled with high-throughput cDNA sequencing have revealed that regions in transcriptome can form highly organized secondary structures [146]. A nascent pre-mRNA synthesized by RNA polymerase-II undergoes co-transcriptional folding, splicing, polyadenylation, and then bound by nuclear RNA-binding proteins for export of mature mRNAs to the cytoplasm. Therefore, it is likely that transcriptome RNA–RNA dynamics must occur during the process. The example of *cis*-acting RNA–RNA base pairing for controlling alternative splicing is found in *Drosophila* DSCAM gene. The exon 6 cluster of DSCAM gene contains 48 alternative exons that are used in a mutually exclusive manner to produce >36,000 distinct spliced mRNA isoforms. The size almost three times more than the number of genes present in fruit fly [147]. Evidence suggests that RNA–RNA base pairing between the docking and selector sites present within the exon 6 cluster dictates which of the 48 alternative exons are used to make the mature DSCAM mRNA [148].

The trans-acting RNA-RNA interaction is known for snoRNA HBII-52 which regulate alternative splicing of the serotonin receptor 2C pre-mRNA. Here, an 18-nucleotide complementary RNA region between the snoRNA and a splicing silencer element in the serotonin receptor pre-RNA leads to alternative exon usage [149,150].

The interaction between chromatin binding proteins and splicing factors can also influence splicing patterns. For example, chromatin-binding protein CHD1 contains two tandem chromodomains and binds tightly to covalently modified histone H3 (H3K4me3). It also forms a stable complex with U2 snRNP thus it bridges the spliceosome with histone and regulate splicing [151]. Splicing repressor protein PTB/hnRNPI bind to the histone-binding adapter protein MRG15 where perturbations of MRG15 alter the splicing of the FGFR2 gene

[152]. Histone modifications may alter chromatin states and change RNA polymerase II transcription rates. Biochemical studies have indicated that RNA polymerase II has direct effects on pre-mRNA splicing and polyadenylation [131].

1.7.5 Role of DUBs in splicing regulation

Previously the interaction landscapes studies of mammalian DUBs have highlighted their distribution of interactors linked to several cellular processes [153]. In this study USP4, USP15, and USP39 shared 14% of their interacting proteins linked to mRNA processing [153]. USP39 was found to associate with BCDIN3, a 7SK small nuclear RNA (snRNA) methylphosphate capping enzyme, known to interact with U5/U6-snRNPs [154]. Additionally, USP39 complexes contained 11 known subunits of U5/U6-snRNPs, suggesting the possible role for USP39 in pre-mRNA splicing [153].

In contrast, USP4 and USP15 associated with five subunits of the U5/U6-snRNP (PRPF31, PRPF3, BCDIN3, PRPF4, and PPIH/cyclophilin H) and subunits of the LSM mRNA binding complex (LSM2-8) [153]. Interestingly, both USP4 and USP15 associated with Terminal Uridylyl Transferase (TUT1), previously implicated in 3' uridylation of U6 snRNA [155]. However, TUT1 is not known to associate with the U5/U6-snRNP complex. In reciprocal LC-MS/MS experiments, SART3 complexes contained the majority of U5/U6-snRNP components as well as USP39 and USP4. The PRPF4 complexes contained USP39 and USP15.

Previous studies have revealed the involvement of ubiquitin in pre-mRNA splicing pathways via the U5/U6-snRNP complex in budding yeast [156]. The spliceosomal Prp19 complex/NTC promotes the modification of the U4 snRNP component Prp3 and PRP31 with K63-linked ubiquitin chains. The ubiquitination of Prp3 and Prp31 promotes association with the U5 component Prp8, which allows for the stabilization of the U4/U6.U5 snRNP complex. The DUB USP15 and its substrate targeting factor SART3 (U6 snRNP-binding protein) forms a complex with USP4, and this ternary complex serves as a platform to deubiquitinate PRP3 and PRP31. The ubiquitination and deubiquitination status of these splicing factors regulate rearrangements of the U4/U6.U5 tri snRNP complex during splicing [116,157]. Differences in the associated proteins suggest that these DUBs play distinct roles in ubiquitin-dependent control of RNA splicing and/or decay [153].

1.7.6 The significance of splicing in human diseases

Splicing is a critical regulatory step for gene expression where any abnormality may develop into a human disease condition. Mutations which lead to loss of spliceosomal function, affecting many gene targets or mutations which disrupt the splicing of specific genes are the primary cause of these disorders [158,159]. Here are some examples of diseases that develop due to disruption of spliceosomal biogenesis and/or function.

a) Retinitis pigmentosa

It is an inherited degenerative eye disease that causes severe vision impairment and blindness. Mutations in core spliceosomal proteins like pre-mRNA-splicing factor 3 (PRPF3), PRPF8, PRPF31, phosphatidic acid phosphohydrolase 1 (PAP1), and the human homologs of pre-mRNA-processing 8 (Prp8) and Brr2 are found to be associated with autosomal-dominant retinitis pigmentosa [39,160–162]. This evidence suggests that human retinal cells are sensitive to splicing defects.

b) Spinal muscular atrophy (SMA)

A recessive neuromuscular disease is caused by reduced levels of the survival motor neuron (SMN) protein or homozygous deletion of the *SMN1* gene in humans. SMN protein functions in the biogenesis of spliceosomal small nuclear ribonucleoproteins (snRNPs) and its complete inactivation is embryonic lethal in mouse [163]. A single point mutation in SMN2 gene causes exon skipping and results in a truncated and unstable protein product [164]. Although splicing defects are a predominant feature of severe SMA, they are detectable only relatively late in the disease course, as studied in a mouse model [165] However, what remains unclear is how the partial loss of SMN function causes a neuromuscular disease and requires further investigation.

c) Chronic lymphocytic leukaemia and myelodysplasia

The core components of the U2 snRNP, such as splicing factor 3B subunit 1 (SF3B1) and U2 auxiliary factor 35 (U2AF35), are frequently mutated in these cancers [166,167]. Such mutations might result in defective snRNP assembly, misregulated alternative splicing or accumulation of unspliced mRNA, and may change expression patterns of multiple genes [168]. In addition to genetic mutations, the misregulation of splicing factor levels has often been found associated with various neoplasias [167]. Therefore an extensive change of alternative splicing observed for thousands of gene samples of cancer patients might be

linked to variations in the expression levels of major splicing factors in cancers. Therefore, targeting the spliceosome function may provide a new route for cancer therapy.

d) Mutations in the minor U4atac snRNA results in a rare genetic defect that causes severe growth retardation and infant death. The condition is known as microcephalic osteodysplastic primordial dwarfism (MOPD) type I [169].

1.8 Brief introduction of the project

UBLs are reported to co-purify with splicing complexes thereby can also function as regulators of spliceosomes [170,171]. Several splicing factors have been reported to be ubiquitinated in high-throughput proteomics studies [8,172]. For example, Brr2 ATPase promotes tri-snRNP complex disassembly by catalyzing the unwinding of U4 and U6 snRNAs. Here, the ubiquitination of U5 snRNP core splicing factor Prp8 [173] within the U4/U6-U5 tri-snRNP complex suppresses Brr2-catalyzed disassembly and thus regulate the spliceosomal dynamics [156]. Prp19, the core splicing factor which is a component of nineteen complex, has E3 ligase U-box domain and have ubiquitination activity *in vitro* [115,174]. The modification of Prp3 by SUMO at K289 and K559 is crucial for its recruitment into the spliceosome and efficient pre-mRNA splicing. [175]. SR protein SRSF1 is known to regulate sumoylation by interacting with SUMO-conjugating enzyme Ubc9 and SUMO E3 ligase PIAS1 (protein inhibitor of activated STAT-1) [176]. Many spliceosomal proteins including hnRNPs and SR proteins are reported to be SUMO substrates in human cell lines [177,178].

As mentioned earlier in section 1.6.8, the UBL Hub1 plays an important role in alternative RNA splicing [98]. Hub1 binds to the HIND domain-containing splicing factors Snu66 and/or Prp38 and functions in pre-mRNA splicing by promoting the usage of non-canonical 5′ splice sites in introns. Hub1 thereby promotes alternative splicing of the *SRC1* pre-mRNA in *S. cerevisiae. HUB1* is a non-essential gene in *S. cerevisiae,* its orthologs in *S. pombe* and mammalian cells are essential for viability, perhaps because of the increased prevalence of introns and alternative splicing in *S. pombe* and humans [98,101].

My thesis supervisor, Dr. Shravan Kumar Mishra, worked on Hub1 in Stefan Jentsch's laboratory at Max-Planck Institute of Biochemistry, Martinsried, Germany. In his postdoctoral work, a high throughput screen called epigenetic microarray profiling was carried out to search for genetic interactors of Hub1 in *S. pombe*. The screen was aimed to identify spliceosomal regulators in an intron prevalent organism. In this screen, *S. pombe*

hub1-1 temperature-sensitive mutant [97] was combined with the haploid deletion library of non-essential genes [179]. Since *hub1* plays a specific role in pre-mRNA splicing, many splicing factors such as *saf4, cwf16, cwf18, cwf19, bis1, cwf12, smd3,* and *spf38* showed synthetic sickness with the *hub1* mutant (Fig.1.6 A). Among them, the deletion of *sde2* gene (silencing defective 2) was also synthetically sick with *hub1-1* (indicated by the red colored arrow in Fig.1.8 B).

A – List of Hub1 genetic interactors

Gene	Function in pre-mRNA splicing
saf4	Splicing associated factor, U2 snRNP complex
cwf18	Cdc5-associated complex
cwf19	Cdc5-associated complex
cwf12	Prp19-complex, Second transesterification factor
smd3	Sm snRNP core protein
bis1	Splicing factor, Associated with Prp19-complex
sde2	Silencing defective -2
cwf16	Prp19-complex associated splicing factor
spf38	U5 snRNP component

B

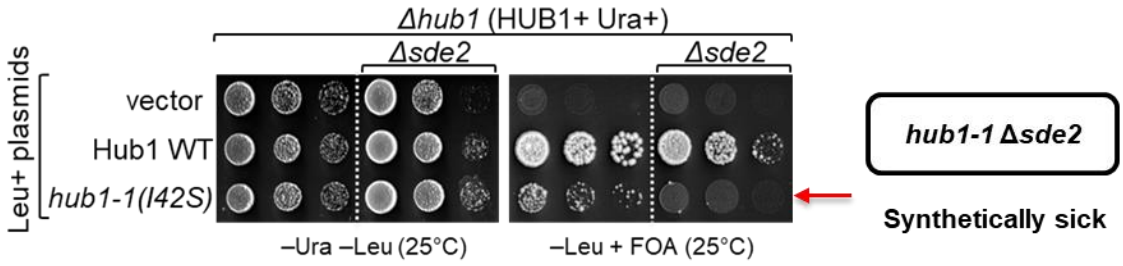


Figure 1.8 - Sde2 genetically interacts with ubiquitin-like protein Hub1.

A) Genetic interactors of *hub1*. *hub1-1* is the temperature sensitive *hub1(I42S)* mutant reported by Yashiroda and Tanaka [97]. A genetic screen was performed with *hub1-1* and the haploid deletion library of non-essential genes in *S. pombe*. *hub1-1* was synthetically sick with the deletion mutants of given genes including *sde2*. Among the top hits, genes with relevance to pre-mRNA splicing are shown.

B) Confirmation of the negative genetic interaction between *hub1-1* and *Δsde2* mutants. *sde2* gene, from ATG to the stop codon, was deleted in *S. pombe Δhub1* strain, which was kept viable with a *ura4+* marked *hub1* expression plasmid with its own promoter and terminator. The resultant strain was transformed with *leu2+* marked expression plasmids expressing wild type (WT) *hub1* or the *hub1-1* mutant. Five-fold serial dilutions of cells from these transformations were spotted on indicated agar plates. Plates were incubated at 25°C. 5' fluoroorotic acid (FOA) (1g/L of media) was used to shuffle-out *ura4+* plasmid. Thus, cells growing on –Leu + FOA plates will have *hub1* mutants in *Δsde2* background.

According to the previous reports from high-throughput studies [180,181], *S. pombe Δsde2* strain showed slow growth under standard conditions. The growth defect was more pronounced at elevated temperatures and under genotoxic stress conditions like in presence of cadmium (generates oxidative stress), valproic acid (HDAC inhibitor), sodium butyrate (HDAC inhibitor), bleomycin (DNA damaging agent), high/low temperature, and thiabendazole (microtubule destabilizing agent) [180–183]. The deletion of *sde2* gene also showed defects in splicing of cytoskeleton constituents and centromeric outer repeat transcripts [182,184].

1.8.1 SDE2 protein

S. pombe Sde2 (abbreviated for silencing defective 2 in *S. pombe*) is a 263 aa long protein with molecular weight of 29.25 kDa. The protein structure prediction program i-TASSER [185] predicted the presence of a ubiquitin fold at the N-terminus and a C-terminal helical domain in Sde2 (henceforth referred to as $Sde2_{UBL}$ and Sde2-C, respectively). An invariant signature motif, GGKGG, separates the moderately conserved $Sde2_{UBL}$ and Sde2-C (Fig. 1.8.1). This motif flanks $Sde2_{UBL}$ and resembles the di-glycine (GG) motif typical of UBL precursors processed by the UBL-specific proteases.

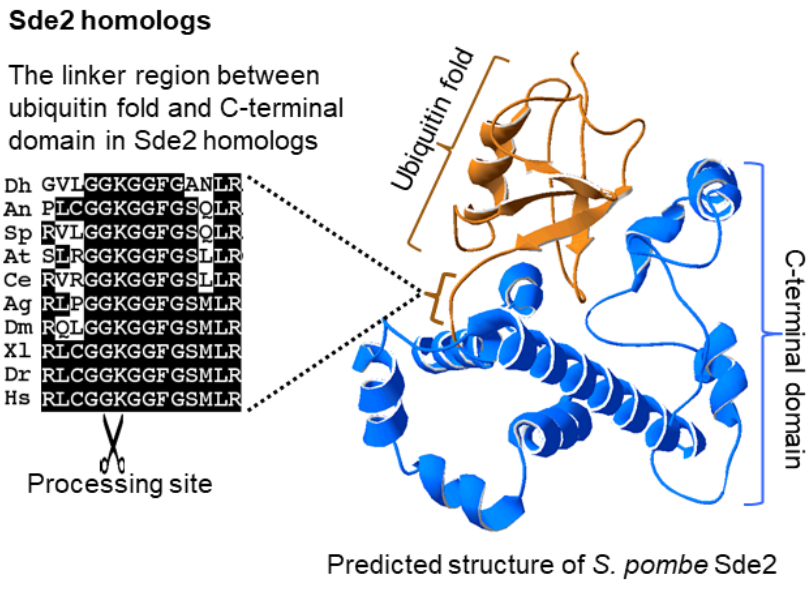


Figure 1.8.1 - Predicted structure of *S. pombe* Sde2 protein from i-TASSER
The structure shows the presence of an N-terminal ubiquitin fold ($Sde2_{UBL}$) and a helical C-terminal domain (Sde2-C). The alignment on the left shows the linker spanning conserved GGKGGFG sequence having the di-glycine motif (GG) between $Sde2_{UBL}$ and Sde2-C in Sde2 orthologs. Dh, *Debaryomyces hansenii* (XP_458854); An, *Aspergillus niger* (XP_001391007); Sp, *Schizosaccharomyces pombe* (NP_594019); At, *Arabidopsis thaliana* (NP_192009); Ce,

Caenorhabditis elegans (NP_506378); Ag,*Anopheles gambiae* (XP_321833); Dm, *Drosophila melanogaster* (NP_651207); Xl, *Xenopus laevis* (NP_001084858); Dr, *Danio rario* (AAH93198); Hs, *Homo sapiens* (NP_689821) **(**data is from Shravan Kumar Mishra).

Putative orthologs of Sde2 exists in organisms from yeast to humans (Fig.1.8.2). However, in the Fungi subphylum Saccharomycotina, a Sde2-like protein is observed in *Debaryomyces* but is absent in *S. cerevisiae, C. albicans,* and *P. pastoris*. Multiple sequence alignment of putative Sde2 orthologs shows that N- and C-termini of Sde2 protein have very low sequence similarity across eukaryotic species. Interestingly, however, a stretch of sequence GGKGGFG is highly conserved between the N and C-termini which harbours two conserved di-glycine motifs (GG) (indicated by the orange line in Figure 1.6.2).

Sde2 was reported to work in telomeric silencing and genomic stability [182]. The human ortholog of Sde2, C1orf55, was present in spliceosomal preparations [111,186], and recently the *S. pombe* protein was also shown to associate with splicing factors [184,187]. Human Sde2 protein was shown to interact with proliferating cell nuclear antigen (PCNA) via a PCNA-interacting peptide (PIP) box. Mono-ubiquitination of PCNA was shown to combat UV induced replication lesions by recruiting DNA damage repair factors to the lesions. Ubiquitin E3 ligase $CRL4^{CDT2}$ dependent degradation of PCNA-associated Sde2 was shown to mediate proper cell cycle progression [188].

We initiated our work after knowing that Sde2 is a genetic interactor of Hub1 in *S. pombe* and it interacts with spliceosome. The predicted ubiquitin fold and presence of a conserved diglycine motif in Sde2 tempted us to hypothesize that, Sde2 might be working like Hub1 in pre-mRNA splicing as a splicing regulator. We got interested in knowing Sde2's function in splicing and its mode of regulation in *S. pombe*.

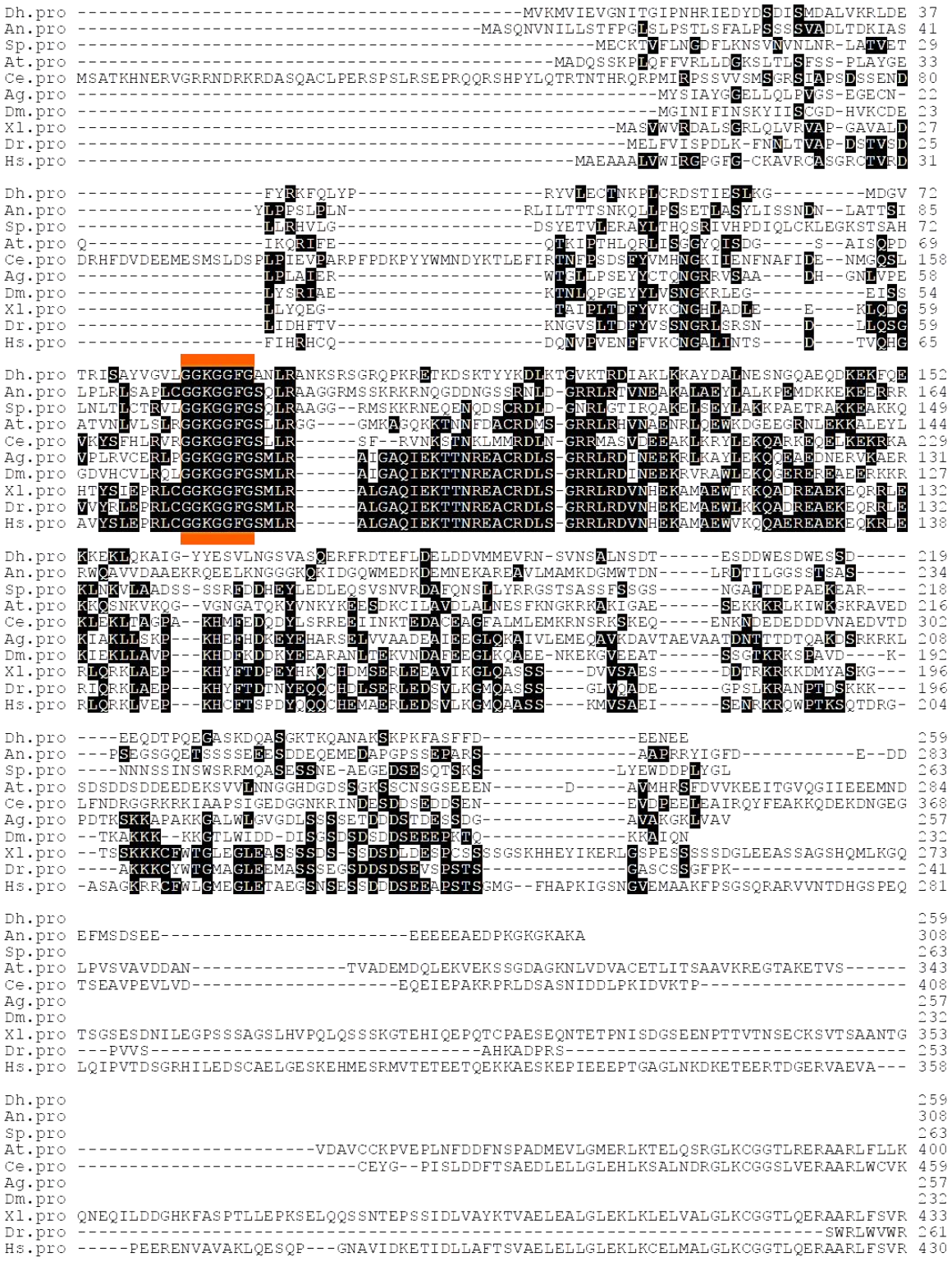


Figure 1.8.2 - Sde2 is conserved from yeast to humans

Alignment of putative Sde2 protein orthologs from different eukaryotes. Abbreviations used are (respective NCBI protein accession numbers are given in parentheses): Dh, *Debaryomyces hansenii* (XP_458854); An, *Aspergillus niger* (XP_001391007); Sp, *Schizosaccharomyces pombe* (NP_594019); At, *Arabidopsis thaliana* (NP_192009); Ce, *Caenorhabditis elegans* (NP_506378); Ag, *Anopheles gambiae* (XP_321833); Dm, *Drosophila melanogaster* (NP_651207); Xl, *Xenopus laevis* (NP_001084858); Dr, *Danio rario* (AAH93198); Hs, *Homo sapiens* (NP_689821). (Data is from Shravan Kumar Mishra)

1.9 Objectives of the study

1. *Whether Sde2 is cleaved like ubiquitin precursors and whether the processing is required for its function*?

The presence of a ubiquitin-fold with a conserved di-glycine (GG) motif is a signature of the UBL family proteins [89]. The predicted protein structure of Sde2 showed the presence of a conserved di-glycine motif (GG) flanking the N-terminal ubiquitin-fold (Fig.1.6). Most of the UBL family proteins are known to get cleaved after the di-glycine motif, an essential process for their activation and function. In correlation to this fact, we speculated that Sde2 might get processed similarly to other ubiquitin-like protein in *S. pombe*.

2. *If Sde2 indeed got processed, then what are its processing enzymes and what is the role of* $Sde2_{UBL}$?

Ubiquitin/UBLs are known to be processed by deubiquitinating enzymes or UBL-specific proteases. These proteases help in activation of UBL precursors and also mediate the removal of Ub/UBLs from the conjugated substrates. Therefore we asked whether any protease is involved in processing of Sde2 and if it indeed got processed then what would be the fate of the cleaved N-terminal $Sde2_{UBL}$ and Sde2-C.

3. *What is the mode of Sde2's spliceosomal association*?

S. pombe Sde2 was shown to associate with splicing factors [184,187], and the finding was confirmed by coimmunoprecipitation experiments and proteomic studies in our lab. Therefore we wanted to know its mode of spliceosomal association and function.

Materials and Methods

2.1 Materials

2.1.1 Chemicals and plastic wares

All chemicals used in the study were either of analytical or molecular biology grades and were obtained from commercial sources. Media components, fine chemicals, and reagents were purchased from Sigma Aldrich, USA, HiMedia, India, Merck Ltd. USA, Difco USA, and Formedium, UK. All plastic wares used for molecular biological and bacteriological works were purchased from Abdos labtech, India, and Tarsons, India.

2.1.2 Molecular biology reagents

Restriction enzymes, T4 DNA ligase, Alkaline Phosphatase (CIP), Phusion DNA polymerase, Taq DNA polymerase, Vent DNA polymerase and other modifying enzymes, their buffers, dNTPs, DNA and protein molecular weight markers were purchased from New England Biolabs, Invitrogen, Sigma Aldrich, USA. Gel-extraction and plasmid miniprep kits were obtained from Qiagen, USA or Bioneer, Korea. RNA isolation kits were procured from, Invitrogen, USA.

2.1.3 Antibodies and antibody-coupled beads

The antibodies anti-Myc (polyclonal), anti-FLAG M2 (Clone M2), anti-haemagglutinin (HA, clone HA-7), anti-HA (polyclonal), anti-rabbit-HRP and anti-mouse-HRP (raised in Goat) were obtained from Sigma Aldrich. Antibody-coupled beads of Anti-HA-rabbit (H6908), anti-MYC (A7470) and anti-FLAG (A2220) were also from the same source. Polyclonal antibody against recombinant purified GST-tagged Sde2-C was raised in rabbit (MERCK), and coupled to protein A/G beads (Pierce) for immunoprecipitation experiments. All primary antibodies were used at 1:5000 dilution in PBST and secondary antibodies were used at 1:10000 dilution in PBST with 5% skimmed milk. For detection of endogenous and recombinant Sde2, anti-Sde2 antibody was used at 1:1000 and 1:5000 dilution respectively in PBST. Around 20-25 µl of antibody coupled beads per 100 OD cells were used for immunoprecipitation experiments.

2.1.4 Media

All the media, buffers and stock solutions were prepared using Millipore distilled water and were sterilized, as recommended, either by autoclaving at 15 lb/inch^2 (psi) pressures at 121ºC for 15 minutes or by using membrane filters (HiMedia, India) of pore size 0.2-0.45 µm (for heat-labile compounds).

- *Luria-Bertani (LB) medium*: 10 g tryptone, 5 g yeast extract and 10 g NaCl were dissolved in 800 ml of distilled water. The pH was adjusted to 7 and the volume was made up to 1 liter and autoclaved.

- *Luria-Bertani (LB) Agar*: 10 g tryptone, 5 g yeast extract and 10 g NaCl were dissolved in 800 ml of distilled water. The pH was adjusted to 7 and the volume was made up to 1 liter and autoclaved. Desired antibiotics (ampicillin (100 µg/ml) or kanamycin (50 µg/ml) were added as per requirements.

- *Yeast-extract with supplements (YES) medium*: Per liter: 5 g yeast extract, 2 g casamino acids, 30 g glucose, 0.1 g adenine, 0.1 g uridine, 0.1 g leucine, 0.1 g histidine, 20 g agar; when selecting for the kan-MX4 marker using G418/geneticin resistance (Sigma G-5013), G418 plates were made by dissolving 200 mg/l. For nat-NT2 cassette, Nat plates with 75 µg/ml and for Hph-NT1, hygromycin plates with 100 µg/ml

- *Synthetic defined (SD) media:* Per liter: 6.7 g yeast nitrogen base, 2 g required supplements dropout mixtures for auxotrophies (e.g. leucine adenine, uracil) are added as required, 20 g glucose, and for making plates 20 g agar was used. Expression constructs under the *nmt81* promoter were induced in the absence of thiamine, and 5 µg/ml thiamine was used to repress the promoter.

- *Edinburgh minimal medium (EMM) medium:* Per liter: 3 g potassium hydrogen phthalate, 2.2 g Na2HpO4, 5 g ammonium chloride, 2 g required supplements mixtures for auxotrophies (e.g. leucine adenine, uracil) are added as required, 20 g glucose, 20 ml of 50x salt stock, 2 ml of 500x vitamins and minerals stock, and for plates 20 g agar was used. Expression constructs under the *nmt81* promoter were

induced in the absence of thiamine, and 5 µg/ml thiamine was used to repress the promoter.

- *SOB medium*: Per liter: 20 g bactotrypton, 5 g yeast extract, 0.5 g NaCl are dissolve in 1 liter water. Then adjust the pH to 7.5 and autoclave. After sterilization add 20 ml 0.5 M $MgCl_2$ (sterile)

2.1.5 Buffers and stock solutions

- *Cell lysis buffer* (for genomic DNA isolation of yeast): 2% Triton X-100, 1% SDS, 100 mM NaCl, 10 mM Tris-Cl (pH-8.0), and 1 mM EDTA.
- *10x Tris-EDTA (TE) buffer (pH 8.0)*: 10 mM Tris-HCl (pH 8.0), and 1 mM EDTA
- *50x TAE:* Per liter: 242 g Tris base, 57. 1 ml glacial acetic acid, and 100 ml of 500 mM EDTA (pH 8.0). Autoclaved and stored at room temperature
- *High Urea (HU) buffer:* 8M urea, 5% SDS, 200 mm Tris pH 6·8, 1 mM EDTA, with bromophenol blue, and 1.5% dithiothreitol (DTT) was added before use.
- *30% Acrylamide Mixutre:* 29.2% acrylamide, and 0.8% methylbisacrylamide
- *Resolving Gel (12%):* 1.25 ml 1.5 M Tris-HCl (pH 8.8), 50 µl 10% SDS, 2 ml 30% acrylamide solution, 50 µl 10% ammonium persulphate (APS), 5 µl Tetramethylethylenediamine (TEMED), and 1.7 ml water.
- *Stacking Gel (4%):* 0.5 ml 0.5 M Tris-HCl (pH 6.8), 20 µl 10% SDS, 0.26 ml 30% acrylamide solution, 40 µl 10% APS, 4 µl TEMED, and 1.2 ml water.
- *10x SDS buffer (pH 8.3):* Per liter: 30 g Tris base, 144 g glycine and 10 g SDS
- *20x MOPS buffer (pH 7.7):* 50 mM MOPS, 50 mM Tris base, 0.1% SDS, and 1 mM EDTA.
- *10x Semi-dry transfer buffer:* Per liter: 29.3 g Glycine, 58.2 g Tris base, and 4 g SDS. For transfer 1x buffer with 5% methanol was used.
- *10x Tris-Buffered Saline (TBS) buffer (pH 7.6):* Per liter: 24.2 g Tris base, and 80 g NaCl. For washing, 1x TBS with 0.1% tween 20 was used.

- **For yeast competent cells preparation**
  - *SORB (pH 8):* 100 mM lithium acetate, 10 mM Tris–HCl (pH 8), 1 mM EDTA (pH 8), 1M sorbitol, filter sterilized and stored at room temperature.
  - *40% PEG:* 100 mM lithium acetate, 10 mM Tris–HCl (pH 8), 1 mM EDTA (pH 8), 40% PEG, filter sterilized and stored at 4°C.
  - *Salmon sperm DNA (10 mg/ml):* It was denatured at 100°C for 10 minutes, cooled on ice and stored at -20°C. For yeast competent cell preparation, 40 µl of denatured salmon sperm DNA was used per 50 ml of culture.

- **For E. coli competent cells preparation**
  - *Transformation buffer - 1 (100 ml)*

Stock	**Final concentration**	**Vol. from stock**
1 M MOPS (pH 6.5 with KOH)	10 mM	1 ml
1 M KCl	100 mM	10 ml
1M $MnCl_2$	45 mM	4.5 ml
1M $CaCl_2$	10 mM	1 ml
1 M KAc (pH 7.5 with HCl)	10 mM	1 ml
Sterile water		82.5 ml

  - *Transformation buffer - 2*
    This buffer is nearly identical to Transformation buffer - 1: The difference is that you add in this case additionally 12.5 ml of 80 % glycerol and instead of 82.5 ml of water just 70 ml (total volume = 100 ml)
  - *Transformation buffer – 3*
    100 CaCl2
    50 M MgCl2

- **For purification of GST-Sde2-C protein**
  - *Lysis buffer* : 100 ml Phosphate Buffered Saline, 5% Glycerol, 1 mM PMSF, 1 tablet of protease inhibitor (Add Lysozyme - 1-4 mg/ml at the time of use)
  - *High salt buffer***:** 500 mM NaCl, 50 mM Tris (pH - 7.5), 1 mM MgSO4, 1 mM EDTA, 1 mM PMSF

- *Elution buffer***:** 100 mM NaCl, 50 mM Tris (pH 8), 10 mM Glutathione

- *10 mM Phosphate Buffered Saline, pH 7.4*: 0.26 g KH2PO4, 2.17 g Na2HPO4-7H20, 8.71 g NaCl, 0.201 g KCl, 800 mL dH20. Adjust pH to 7.4 and bring volume to 1 L with dH20.

- **For coupling of GST-Sde2-C protein**
  - *Coupling/Wash Buffer*: 0.1 M sodium phosphate, 0.15 M NaCl, pH 7.2
  - *Protein or peptide*: Dissolve 1-20 mg protein or 1-2 mg peptide in 2-3 ml of Coupling/Wash Buffer. For proteins already in solution, dilute four-fold in Coupling/Wash Buffer or desalt/dialyze into Coupling/Wash Buffer. Completely remove primary amine-containing buffer (e.g., Tris or glycine).
  - *Quenching Buffer*: 1 M ethanolamine, pH 7.4

- **For affinity purification of antibody**
  - *Binding/Wash Buffer*: Phosphate-buffered saline
  - *Sample*: Prepare antigen or other molecule in Binding/Wash Buffer or dilute sample 1:1 in Binding/Wash Buffer
  - *Elution Buffer*: 0.1-0.2 M glycine-HCl, pH 2.5-3.0
  - *Neutralization Buffer (optional)*: 1 ml of high-ionic strength alkaline buffer (1 M phosphate, pH 9)

- **For immunoprecipitation in yeast**
  - *IP Cell-lysis buffer*: 150 mM NaCl, 5 mM MgCl2, 1 mM EDTA, 50 mM Tris-Cl (pH 7.5), 5-10 % Glycerol. Store the buffer at 4°C. Add to 50 ml before use: 500 µl of 100 mM PMSF (dissolved in isopropanol), 1 tablet of protease inhibitor (Thermo Scientific), 1 tablet of phosphatase inhibitor (Roche), 1% Triton X100
  - *IP Wash buffer 1*: *Cell-lysis buffer* with 1% triton X-100 and 1 mM PMSF
  - *IP Wash buffer 2*: *Cell-lysis buffer*, without protease inhibitor, PMSF and triton X-100

2.2 Methods

2.2.1 Yeast and bacterial strain maintenance

S. pombe strains were grown to stationary phase (around 1.5 OD_{600}) in YEL media at 30°C for 16–18 hours with shaking at 250 × g. The saturated yeast culture is mixed in 1: 0.4 ratios with 50% (v/v) sterile glycerol and after mild vortexing, stored at -80°C. For *S. cerevisiae* strains, YPAD (Yeast extract peptone and dextrose) medium is used and grown at 30°C for 12-14 hours. Rest of the process of glycerol stock preparation is similar to that used for *S. pombe* strains. For experiments, strains were revived from glycerol stocks on above mentioned solid agar media plates and maintained at standard growth conditions. The transformed yeast strains were plated on synthetic complete (SC) medium with supplements as per the requirement to keep the selective pressure on the plasmid. For *E. coli* strains, Luria and Bertani medium was used and cells were grown at 37°C for overnight at 250 × g. Then 500 µl of culture is mixed with an equal volume of sterile 50% glycerol. After mild vortexing, the stocks were stored in cryovials at -80°C. Yeast strains used in this study are listed in Table 2.1.

2.2.2 Yeast genomic DNA isolation

Yeast strains were grown till ~1.5 OD_{600} in 5 ml YEL media at 30°C for 16–18 hours at 250 × g. Cells were harvested by centrifugation at 2000 × g for 5 min at room temperature and washed with sterile distilled water. Cells were lysed using glass beads method in the presence of 200 µl cell lysis buffer and 200 µl phenol:chloroform: isoamyl alcohol [25:24:1 (v/v/v)] by vigorous vortexing for 8 min then transferred to ice for 1 min. After vortexing, cells were suspended in 200 µl TE buffer and centrifuged at 21000 × g for 10 min at room temperature. The supernatant was transferred to a new microcentrifuge tube and mixed with equal volume of chloroform; vortexed and centrifuged for 5 min. The aqueous layer on top was carefully transferred to a fresh tube with 1 ml of 100% ethanol and vortexed for few seconds. The tube was incubated at -20°C for 1 hr and then centrifuged at 18000 × g for 10 min. The supernatant was removed and 400 µl TE buffer with 6 µl of RNase A (5 mg/ml) was added to the pellet and incubated at 37°C for 30 min. For genomic DNA precipitation, 10 µl 4 M ammonium acetate and 1 ml 100% ethanol were added and kept -20°C for 1 hr. The tube was centrifuged at 21000 × g for 10 min. The pellet was washed with 1 ml 70% ethanol and then dried in a vacuum concentrator and suspended in 50 µl TE buffer. The concentration

was measured using nanodrop; usually, genomic DNA with 1 µg/µl concentration was achieved.

2.2.3 Preparation of yeast competent cells

Competent cells preparation for *S. pombe* and *S. cerevisiae* was done as per published protocols [189]. Briefly, *S. pombe* cultures were grown in YEL media at 30°C with shaking for 16-24 hours and then reinoculated in fresh YEL media at OD_{600} of 0.2-0.3. Cells were further incubated at 30°C with shaking at 250 rpm till the OD_{600} of around 0.5–0.7. The cells were harvested by centrifugation at 1800 × g, for 5 min, room temperature and then washed with sterile millique water. Cells were washed once with 0.1X SORB (for 50ml of culture volume use 5 ml of SORB). After centrifugation, SORB was removed and the pellet was resuspended in fresh 360 µl SORB and 40 µl denatured salmon sperm DNA (10 mg/ml stock solution) per 50 ml culture. Aliquots of competent cells were stored at -80°C. For *S. cerevisiae* similar protocol was followed except the YPAD medium is used for growing the cells.

2.2.4 Transformation of *S. pombe* and *S. cerevisiae*

The transformation of *S. pombe* strains was carried out by lithium acetate method [189]. 10 µl of competent cells were mixed with 1 µl-2 µl of plasmids in a sterile microcentrifuge tube and six-fold sterile 40% PEG is added. After vortexing, cells were incubated at 30°C for 30 minutes. Heat shock was given at 42°C for five minutes, and cells were kept on ice for 5 minutes. After the addition of 100 µl sterile water, the whole solution was then plated on selection plates and incubated at 30°C for 3-4 days. For S. cerevisiae transformation identical protocol was used except the heat shock time was given for 20-25 min at 42°C.

2.2.5 Preparation of *Escherichia coli* competent cells

E. coli strain was streaked on a fresh agar plate (LB agar) without antibiotics and incubated overnight at 37°C. A single colony was inoculated in 5 ml of LB medium and incubated overnight at 37°C shaking at 250× g. Then 500 µl of the overnight grown culture was transferred to 50 ml of SOB medium and incubated 37°C till the suspension reached the OD_{600} of 0.5 (after approx. 2.5 h). Then culture was kept on ice for 10 min and then transferred into falcon tubes for centrifugation at 2500 × g for 10 min at 4°C. The supernatant was removed, and the pellet was resuspended carefully by adding 25 ml Transformation

buffer-1. The suspension was incubated for 10 min on ice and again centrifuged for 10 min at 2500 × g, 4°C. After removing supernatant, 4 ml of Transformation buffer-2 and 140 µl of DMSO was added to the pellet followed incubation of suspension on ice for 15 min. After addition of another 140 µl of DMSO, the cells were distributed in aliquots of 200 µl or 400 µl and stored immediately at -80°C (precooled Eppendorf tubes were used). For transformation equal volume of Transformation buffer -3 was added to the cells and 100 µl of the suspension for used for transformation.

2.2.6 Transformation of *E. coli*

100 µl aliquot of competent cells from -80°C refrigerator was taken on the ice and added with an equal volume of transformation buffer – 3 (kept at 4°C). Around 100-200 ng of plasmid was added to 100 µl competent cells and incubated for 30 min on ice. Heat shock was given at 37°C for 60-85 seconds on thermoblock or water bath and the cells were immediately transferred to ice. After 5 min on ice, 1 ml of LB broth was added to the cells and incubated at 37°C for shaking, 1 hr. Around 100 µl of culture from the tube was plated on selective antibiotic containing LB agar plates. The plates were incubated at 37°C for overnight to get the transformants.

2.2.7 Complementation assays and growth phenotype study in yeast

To complement the growth defect phenotypes of *S. pombe* mutants, the competent cells of mutant strains were transformed with respective plasmids and plated on selection plates (SC medium). The plates were incubated at 30°C until the growth of transformants is observed. After growth, the transformed cells were resuspended in sterile water, and OD_{600} is measured. The cells were diluted at five-fold serial dilution in a microtiter plate with an initial OD_{600} of 2.0. The dilution spotting was done on plates with 5 µg/ml thiamine and without thiamine. For uniform spotting, a custom made spotter was used to take diluted cells from microtiter plates to the agar plate. Following spotting, the plates were incubated at optimal and unfavourable temperatures until growth was observed. For study of growth phenotype of *S. pombe* chromosomal mutants, YES media was used and similar protocol was followed as mentioned for deletion strains.

2.2.8 QuikChange site-directed mutagenesis (SDM)

All point mutations, insertions or deletions on plasmids were created by QuikChange site-directed mutagenesis (Agilent) using specific-primers and high-fidelity *Pfu* DNA polymerase.

The primers designed for SDM harbour the desired mutations and are flanked by unmodified nucleotide sequences complementary to the sequence of opposite strands of the plasmid. For SDM, the PCR amplification was done for 18 cycles using 10-50 ng of the template with the annealing at 55°C for 1 min and extension at 68°C for 1-2 min/kb of plasmids length. For plasmid with size ≥ 10 kb, an extension time of 2 min/kb was used. After completion of PCR, the mixture was treated with 1 µl Dpn1 enzyme at 37°C for 1 hour to digest the parent template. Subsequently, 5 µl of amplified product was transformed in 100 µl of DH5α competent cells and plated on the selective antibiotic plate.

2.2.9 Overlap extension (SOE) PCR

To make gene chimeras, overlap extension (SOE) PCR method was used as per the following protocol. The primers were designed upto 30-40 bp in length by keeping the minimum complementary region of around 20 bp. The SOE forward and reverse primers had flanking sequences complementary to each other such that, when PCR products of two different genes are mixed in equal proportion, then, these flanking regions can bind together and DNA polymerase can extend from their 3' ends. The normal forward and reverse primers were used together with SOE primers for individual amplification of the genes. The PCR products were obtained using Phusion polymerase (from NEB). The PCR products were run on agarose gel and purified using DNA gel extraction kit (Bioneer). The concentration of PCR products was measured and used as templates for SOE PCR. Around 40-45 ng of larger PCR fragment was taken and accordingly the amount of the smaller PCR fragment was mixed in 1:1 ratio of DNA molecules in the final SOE PCR reaction. For SOE PCR, Vent or Phusion polymerase were used from NEB.

2.2.10 Protein isolation by trichloroacetic acid (TCA) precipitation

Grow the yeast cells to log phase in desired media and harvest 1 OD_{600} cells by centrifugation at 2400 × g for 5 min. According to a published protocol [189], the pellets were then resuspended in freshly prepared 1 ml of 7.5% β-mercaptoethanol with 2 N NaOH solution. The suspension was vortexed and kept on ice for 10 min. Then 200 µl of 55% trichloroacetic acid (TCA) was added and further incubated on ice for 10-15 min. TCA precipitation was followed by centrifugation at 21000 × g for 15 min at 4˚C. The supernatant was discarded, and tube was centrifuged again and leftover traces of TCA/supernatant were removed with vaccusip. For protein extraction 50 µl HU buffer with 1.5% DTT was added

and heated at 65˚C, 10 min with shaking. The mixture was centrifuge at 18000 × g, 5 min, room temperature; 10 µl of the isolated protein lysates were used for immunoblot assays.

2.2.11 Western blot (WB) assay

For immunoblot assays, 1 OD_{600} cells from exponentially growing culture were harvested. Primary cultures were grown in synthetic defined media till saturation at optimum conditions; then reinoculated to secondary culture in EMM media (without thiamine) and finally to tertiary culture in EMM media to induce protein expression from the thiamine-repressible *nmt81* promoter. Total proteins were isolated by TCA precipitation and 10 µl of the isolated proteins were loaded on SDS-PAGE and transferred to PVDF membrane for two and half hour at 110 mA (per 56 cm^2 PVDF membrane). Western blot semi-dry transfer apparatus from SCIE-PLAS was used for gel to PVDF membrane transfer of proteins. Blocking of membrane was done with 5% skimmed milk for 1 hour at room temperature. The membrane was then incubated with the primary antibody for 2-3 hrs, followed by 5 min washing for 4 times with 1x TBST buffer. After washing, the blot was incubated with HRP-conjugated secondary antibody for 1 hour at room temperature. Blot was washed with 1x TBST buffer, and proteins were visualized using chemiluminescent detection reagent from Pierce or Bio-Rad. The signals were captured using X-ray films and documented.

2.2.12 Co-immunoprecipitation (Co-IP) assay

- *Cells harvesting:* Cells were grown to log phase, around 0.6-0.8 OD_{600} and total 100 OD cells were harvested by centrifugation at 900 × g, 5 min at 4°C. After centrifugation supernatant was discarded and the pellet was resuspended in cell-lysis buffer with PMSF, phosphatase and protease inhibitors, and flash frozen in liquid nitrogen and stored at -80°C. The assay is described previously [98].
- *Cell-lysis:* Total cell lysates were prepared by mechanical grinding of frozen pellets with liquid nitrogen in the presence of IP cell-lysis buffer. The total cell suspension volume for 100 OD cells was kept to 1 ml. Lysates were pre-cleared two times by centrifugation at 10,000 × g for 10 min at 4°C to remove cell debris.
- *Immunoprecipitation (IP):* After pre-clearing, the supernatant is transferred to fresh micro-centrifuge tube and immunoprecipitation was done using appropriate antibody-tagged beads (20-25 µl beads/100 OD_{600} cells) for 3 hours at 4°C on slow speed rotator. After immunoprecipitation unbound proteins were washed away by centrifugation at 900

× g at 4°C for 2 min, first with diluted lysis buffer; then three times with wash buffer 1 and finally by wash buffer 2 (without Triton X-100). The supernatant was discarded thoroughly using vaccusip and both input samples (2%), as well as immunoprecipitated proteins, were extracted by heating at 65°C for 10 min in the presence of 25 µl HU buffer. After centrifuging it at 18,000 × g, 5 min, room temperature; 10 µl of the eluted proteins were loaded on NU-PAGE gel and co-immunoprecipitated proteins were analyzed by immunoblotting with respective antibodies.

2.2.13 Localization of Sde2

For localization of Sde2-C, the gene was chromosomally tagged with EGFP by homologous recombination [190]. For $Sde2_{UBL}$ and precursor Sde2 localization, an EGFP-Sde2 fusion and EGFP-Sde2 (G<u>V</u>KGG) clones were prepared respectively under nmt promoter. *S. pombe* cells transformed with these clones were grown in EMM-leu medium. Cells were harvested at OD_{600} 0.7-0.8 and used for localization studies. For Sde2 chromosomally tagged strain YEL media was used. When the culture reaches to 0.6-0.7 OD_{600} then 1 µl DAPI (1 mg/ml DAPI stock solution) per ml of culture was added and grown for 1 hour before harvesting.

For mounting cells for imaging the protocol is adapted and modified from [191]. Agarose pads were prepared by melting agarose (2 %) in water at 95°C heat-block for 5-10 min and by adding equal volume of 2X EMM-leu or YEL. Around 200 µl of this solution was poured in the middle of glass slide. On both sides of the glass slide, spacers with a thickness of ~0.5 mm were kept and another glass slide was placed on top. The agarose solution will evenly spread between the glass slides. To make spacers, use strips cut out of cardboard, such as that delivered with restriction enzymes. Meanwhile, 1 ml of cultured cells were pelleted at 1,000 × g for 1 min and the supernatant was removed. The pellet was re-suspended in fresh 200 µl medium. After 2-3 min spacers were carefully removed along with the top glass slide. Around 5 µl of cells were spread onto the pad with the help of micropipette and allowed to dry (~1 min) before covering with cover slip. Coverslip was sealed with appropriate sealant and imaging was done at RT. Images were captured with NIKON Ni-E fluorescence microscope equipped with fluorescence optics and Nikon A1 confocal imaging system equipped with apochromat 60x/1 NA oil immersion objective lens. Imaging of bright field was done using Nikon DS-L3 camera attached onto the same microscope. Images were processed with ImageJ software.

2.2.14 Chromosomal tagging of splicing factors

Chromosomal tagging and gene deletion was done following published protocols [190]. Briefly, to monitor the relative abundance of splicing factors, sequence encoding either 6HA, EGFP, or 9MYC epitope tag with Nat-NT2 or KanMX4 or HphNT1 cassette for resistance against Nourseothricin/G418/Hygromycin antibiotic was inserted chromosomally at the C-terminus of selected splicing factors. The cassette was amplified using long primers (100 bp) that contain sequences of homology to the genomic target location with a mixture of 2 units of Taq polymerase and 0.4 units of Vent polymerase. The PCR product was ethanol-precipitated and 10 µl of it was transformed in *S. pombe* strains. After transformation, the strains were revived in YEL media for 12-16 hours with shaking followed by selection on respective antibiotic containing YES agar plates. The transformants were screened for tagging of desired gene by immunoblotting assays using tag specific antibody (Fig.2.2).

2.2.15 Chromosomal mutagenesis of *sde2* gene

A plasmid was made which contained a NotI insert comprising of the 500-bp *sde2* promoter, *sde2* ORF with the coding sequence of 6HA epitope tag, a nat-NT2 cassette for resistance against Nat antibiotic, and 500 bp of the *sde2* terminator. The plasmid was mutagenized by site-directed mutagenesis to obtain *AAKGG* and *GGMGG*, *GGRGG* and *GGTGG* versions of *sde2*. For making *MGG-sde2-C*, the sequence encoding amino acids 2–85 of $sde2_{UBL}$ was deleted by SOE PCR. The NotI digested inserts from these plasmids were purified and transformed in *S. pombe* strains. The revival in YEL media was done for not more than 8-12 hrs. The revived cells were plated on YES-Nat plates. Here the slow growing or smaller colonies were picked up with an assumption that mutation at the desired gene locus has caused a growth defect. The chromosomal mutations were screened by growth phenotypes, immunoblotting assays using anti-HA antibody and/or sequencing of genomic PCR fragments covering the mutated regions (Fig.2.2). On an average, 30-40 colonies were screen by growth phenotype.

2.2.16 RNA isolation and RT-PCR

- *Cells harvesting:* RNA isolation and cDNA synthesis were done as described previously [192]. Briefly, 5 OD600 cells in logarithmically growing phase were harvested at 30°C (untreated control) or after 15 minutes of heat shock at 37°C by filtration and pellets were stored at -80°C after snap freezing with liquid nitrogen.

- *RNA isolation:* Total RNA was isolated by hot acid phenol method using 15 ml phase lock gel tubes. Briefly, pellets were resuspended in acid phenol: chloroform and AES buffer [50 mM NaAcetate (pH 5.3), 10 mM EDTA, 1% SDS] by vortexing. Then the pellets were transferred to 65°C water bath for 7-10 min and vortexed thoroughly once every minute. After lysis, cells were incubated on ice for 5 min and entire organic and aqueous phase was transferred to pre-spun 15 ml phase lock gel tubes. The tubes were centrifuged at 1800 × g at 4°C for 5 min. Then, Phenol:Chloroform:IAA (25:24:1) was added to the gel tubes, and again centrifugation was done. Subsequently, chloroform was added to the supernatant, and after centrifugation, the aqueous phase was transferred into a new 15 ml conical tube with isopropanol and 3 M sodium acetate. The conical tubes were vortexed thoroughly and 2 ml isopropanol slurry was centrifuged at maximum speed for 20 min at 4°C. After centrifugation, the supernatant was discarded, and RNA pellets were washed two times with 70% ethanol. The RNA pellets were dried in a vacuum concentrator and finally resuspended in nuclease-free water. It was followed by DNase I treatment for 15 min at room temperature and Zymo-Spin II column was used for clean-up of RNA. Total RNA was quantified with a spectrophotometer by measurement at OD260/280nm (Nanodrop).

- *cDNA synthesis and RT-PCR:* cDNA synthesis from 3µg total RNA was done using RT and random-hexamer primer at 42°C for 16 hours followed by RT-PCR using target-specific primers and products were analyzed by agarose gel electrophoresis. Primers used in RT-PCR assays of splicing targets are listed in table 2.3.

2.2.17 Real-time quantitative PCR (qPCR)

Real-time qPCR reactions were carried out using SYBR green dye-based reagents from Roche in triplicates. *act1* was used as a control and relative expression of the gene of interest was normalized with respect to *act1* expression in wild-type. Primers used in these assays are listed in table 2.3.

2.2.18 GST tagged protein purification protocol

GST-Sde2-C fusion cloned in pGEX-5X1 vector was transformed in *E. coli* BL21 DE3 and plated onto LB-amp plates. The transformants were inoculated in 5 ml LB broth with ampicillin and grown overnight at 37°C. Secondary culture was inoculated in 500 ml of LB-amp broth at initial OD_{600} of 0.05. The culture was kept at 37°C for shaking till it reaches

0.5 OD_{600}. For protein induction, 100 µM IPTG (Isopropyl β-D-1-thiogalactopyranoside) was added to the growing culture and further incubated at 16°C for shanking overnight. The culture was harvested by centrifugation at 5000 × g for 5 min and the pellet was washed once with PBS (Phosphate buffer saline) and flash frozen in liquid nitrogen and stored at -80^{o} C. The pellet was taken out from -80^{o}C and kept on ice. For cell lysis, around 40 ml of Lysis buffer (with 60 µl of 1mM PMSF, 200 µl Protease inhibitor cocktail solution, 0.1 µl/ml Nuclease) was added and mixed properly. The mixture was subjected to sonication for 10 min (10-15 Amplitude, 10 Sec on sonication & 10 Sec cooling) for rupturing the cells. The content was centrifuged at 5000 × g for 10 min at 4^{o}C and the supernatant was carefully transferred to 50 ml oakridge tubes. The oakridge tubes were centrifuged at 70000 × g for 30 min at 4^{o}C and the supernatant was transferred to 50 ml Falcon tubes containing 2 ml of washed glutathione beads (Note: Wash the glutathione beads twice with 10 ml of lysis buffer before addition of supernatant, the beads can be washed in a 50 ml falcon tube). Binding of GST tagged proteins with beads was done for 60 min on ice with shaking (avoid vigorous mixing).

The supernatant was loaded onto the column (3 ml capacity) & allowed to flow through completely. The flow-through was collected in 50 ml falcon tube and passed again through the column (Note: Never let the beads dry). The column was washed with 10 ml of lysis buffer followed by 10 ml of high salt buffer (with 300 mM NaCl). Finally, the column was washed with 10 ml of detergent buffer (with 1% TX-100) and then with 10 ml of PBS.

For protein elution, 500 µl of elution buffer was added to the column and the eluate was collected in 1.5 ml tubes kept on ice. Around 6-7 of such fractions of eluate were collected and the protein concentration in each fraction was measured using Bradford reagent (SIGMA). The fractions containing protein were pulled into one tube and subjected to dialysis. Dialysis was carried out against against PBS (0.1 mM DTT + 0.2 mM PMSF) for 3-4 hours and then against PBS (0.1 mM DTT + 0.2 mM PMSF + 10% glycerol) for overnight at 4^{o}C. Then 500 µl aliquots of purified protein sample were prepared and flash frozen in liquid nitrogen and store at -80^{o} C.

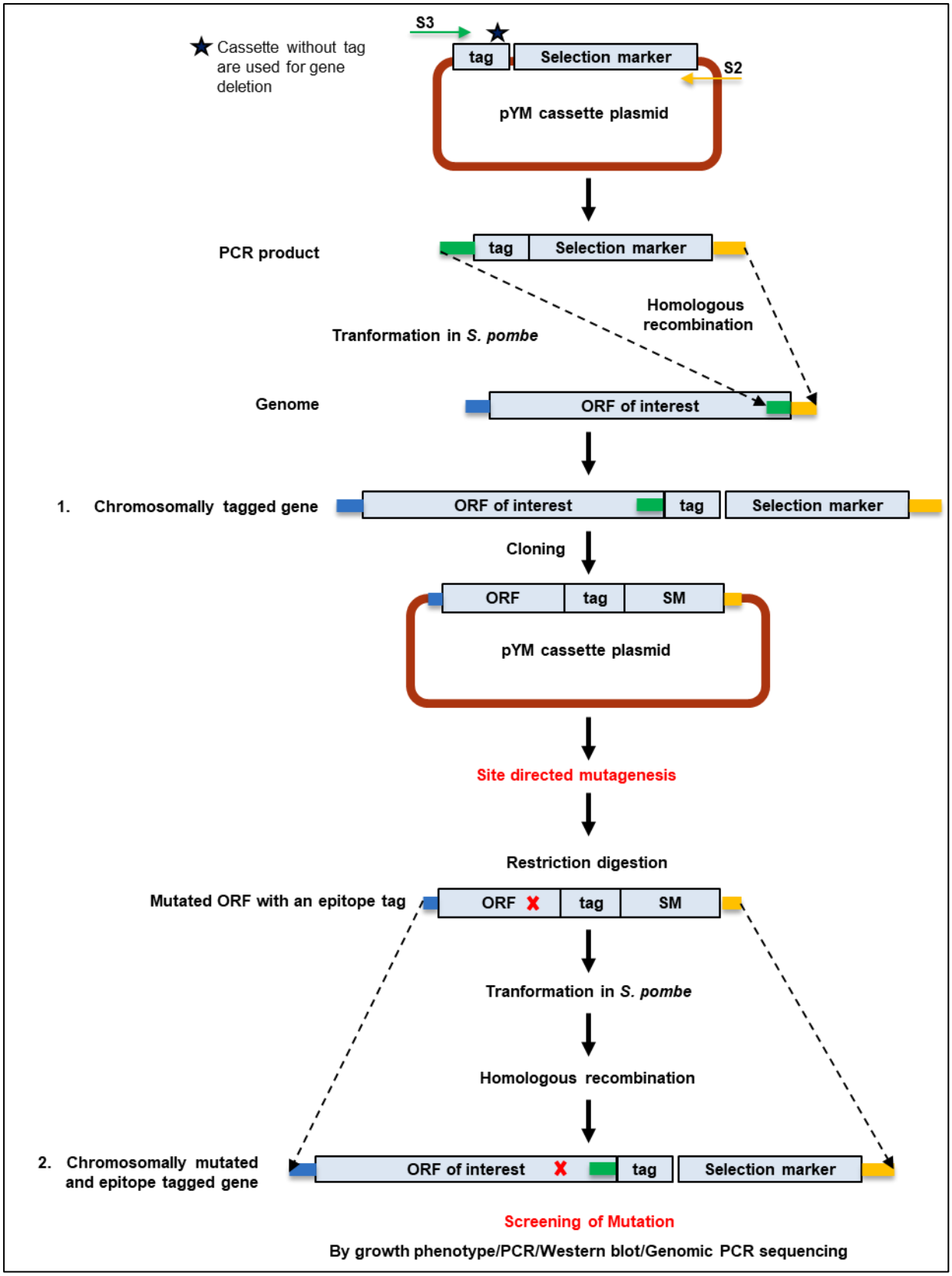


Figure 2.2 - The principle of PCR-based epitope tagging and chromosomal mutagenesis
Schematic illustration of the principle of genomic manipulation of yeast strains using PCR-based strategies. The plasmid contains a cassette, which consists of a selection marker, and additional sequences, which can be promoter sequences and/or sequences that encode for a tag (e.g. GFP). The S1, S2 primers allow amplification of cassettes and targeting of the respective PCR product to

the desired genomic location, which becomes defined by the overhangs provided by the S1- S2 primers (the same colours indicate homologous sequences). Depending on whether a gene deletion or C-terminal tagging **(1)** is to be performed, the respective plasmids are used to amplify antibiotic cassette with or without a tag. Upon transformation, an integration of the cassettes into the yeast genome occurs due to homologous recombination. For chromosomal mutagenesis **(2)**, the desired gene ORF with a C-terminal tag and antibiotic selection marker is cloned in plasmid. Mutation in the ORF is done by site directed mutagenesis and the cassette with overhangs (homologous sequence) is digested, purified and transformed in cells.

2.2.19 Anti-Sde2 antibody purification

a) Coupling Procedure

The agarose slurry was mixed properly and using a wide-orifice or cut pipette tip, 2 ml of the NHS-Activated Agarose Resin Slurry was taken into 15 ml falcon tube. The storage solution was removed by centrifuge at 1,000 × g for 1 minute. The resins were washed with 2 ml of ultrapure water and then 2 ml of coupling/wash buffer. (Note: Do not allow resin to become dry at any time during the procedure.)

1 ml of GST-Sde2-C purified protein (4 mg) was added with 2 ml coupling/wash buffer and total 3 ml of this protein solution was mixed with the washed resins. The reaction was kept for binding with end-over-end shanking for 3 hours at 4^{o}C. Then the mixture was transferred to 5 ml pierce gravity flow through column and allowed to pass through. The column was washed with 1 ml of Coupling/Wash buffer for two times.

To this column, 2 ml of Quenching Buffer was added and both bottom and top caps were replaced. The resins were mixed end-over-end for 15-20 minutes at room temperature. First the top cap was removed and then the bottom cap. The column was kept in a new collection tube and allowed to flow through. The column was washed with at least 6-10 ml of Coupling/Wash Buffer.

b) Affinity purification

The above column was equilibrated by adding 1 ml of Binding/Wash Buffer and allowed to flow through. The process was repeated two times. Around 1 ml of antisera was mixed with 1 ml of binding/wash buffer and loaded into column and allowed to enter the resin bed. The top and bottom caps were replaced and the column incubated with end-to-end shanking at 4^{o}C for 6 hours. The column (after removing top and bottom caps) was placed in a new collection tube, and the solution was allowed to flow through. The flow-through can be saved to analyse binding efficiency. The column was washed with 3 ml of Binding/Wash Buffer.

The bound protein/antibody was eluted by applying 8 ml of Elution Buffer. The fractions of 1 ml (or 0.5 ml) were collected in 1.5 ml Eppendorf tubes. The pH of each fraction was adjusted to neutral by adding 50 µl of Neutralization Buffer per 1 ml of collected eluate. The eluted fractions were measured for antibody concentration using the Bradford assay. The fractions of interest containing antibody were pulled together and dialyzed against PBS (0.1 mM DTT + 0.2 mM PMSF) for 3-4 hours and then against PBS (0.1 mM DTT + 0.2 mM PMSF + 10% glycerol) for overnight at 4^{o}C. Final antibody solutions were prepared in 1X PBS and 50% glycerol and 500µl aliquots of purified antibody were stored at -80^{o} C.

2.2.20 Screening for Sde2-specific proteases

S. pombe deletion strains of selected proteases from a haploid deletion library (Bioneer) were transformed with *pREP81x-3MYC–sde2–3FLAG* plasmid. The transformants were grown for 20 h in the absence of thiamine in EMM-Leu media. 1.0 OD_{600} cells were harvested from logarithmically growing cultures and processed for immunoblot analysis.

2.3.1 Processing studies of *S. pombe* Sde2, Ubiquitin and *Hs* C1orf55

Escherichia coli BL21 (DE3) strain was used to monitor processing of Sde2, Ubiquitin-Sde2-C, C1orf55 by *Sp*Ubp5, *Sp*Ubp15, *Hs*USP7, and protease chimeras. For all co-transformations, the substrates coding genes were cloned in *pPROEX-6HIS* plasmid and DUBs were cloned in *pCDFduet-6HIS empty* plasmid. The co-transformed cells were plated on LB + Ampicillin + Streptomycin antibiotic-containing media. Protein expression was induced in logarithmically growing cells using 100 µM IPTG for 16 hrs at 18°C. 1.0 OD_{600} cells each were harvested and lysed for 10 min with 200 µl B-PER reagent (Pierce) containing 1 µl benzonase (Sigma), 1 mM PMSF (Sigma) and 10 µl protease inhibitor cocktail (Pierce) per 1 ml of B-PER reagent. Then 100 µl of HU buffer was added and heated at 65°C for 10 min and centrifuged at 14000 × g for 10 min to pellet cell debris. 10 µl of lysate was used for immunoblot assay. Sde2 and C1orf55 were detected by anti-Sde2 and anti-FLAG immunoblotting. The proteases were detected by anti-His immunoblotting.

Saccharomyces cerevisiae PJ697a strain was used to monitor processing of *S. pombe* Sde2. For co-expression, *p426ADH-3MYC–Spsde2–3FLAG* plasmid was co-transformed with *p424ADH-Scubp15*, *p424ADH-Spubp5*, *p424ADH-Spubp5 (C222S)*, *p424ADH-Spubp15*, *p424ADH-Spubp15 (C239S)* plasmids. Co-transformants were grown in SC-Ura-Trp minimal media at 30°C. 1.0 OD_{600} cells were harvested from logarithmically growing cells, processed by TCA prep and analyzed by anti-FLAG immunoblotting.

Table 2.1- List of *S. pombe* strains used in this study

- *Nat-NT2*: Encodes for an enzyme nourseothricin acetyl transferase which confers resistance against nourseothricin
- *KanMX4*: Yeast selectable marker conferring kanamycin resistance
- *HphNT1*: Encodes for an enzyme aminoglycoside phosphotransferase which confers resistance against hygromycin

Strain	Relevant genotype	Reference
SP1	h- *ade6-M216, leu1, ura4-D18*	Tanaka's lab
SP9	PEM2 h-; *ade6-M210; leu1-32; ura4-D18; mat1_m-cyhS, smt0; rpl42::cyhR (sP56Q)*	Krogan's lab
SP10	PEM2 *hub1-I42S::Nat-NT2*	This study
SP13	JY741 *hub1::aur1R pUR19-hub1+*	Tanaka's lab [97]
SP20	h+ *JY741 Δsde2:: Nat-NT2*	This study
SP23	h+ *JY741 Δsde2::Nat cdc5–6HA::KanMX4*	This study
SP35	h+ *JY741 Δsde2:: Nat-NT2 prp19–6HA::KanMX4*	This study
SP38	h-*JY741 sde2–6HA::Nat-NT2*	This study
SP40	h-*JY741 sde2–EGFP::KanMX4*	This study
SP48	h- *DY12298 Leu1 mts2*	Li-Lin's lab
SP49	h- *DY12299 Leu1 mts3*	Li-Lin's lab
SP52	h+ *Δubp15::KanMX4, Δubp5::Nat-NT2*	This study
SP55	h+ *Δubp5::KanMX4*	BIONEER
SP56	h+ *Δubp15::KanMX4*	BIONEER
SP77	h- *sde2-6HA (AAKGG):Nat-NT2*	This study
SP78	h+ *sde2-6HA:Nat-NT2 cdc5-9MYC:HphNT1*	This study
SP79	h+ *sde2-6HA (AAKGG):Nat-NT2 cdc5-9MYC:HphNT1*	This study
SP82	h+ *sde2-6HA (K85M):Nat-NT2*	This study

SP83	h+ *sde2-6HA (K85M):Nat-NT2 cdc5-9MYC:HphNT1*	This study
SP85	h+ *MGG-sde2-C-6HA:Nat-NT2*	This study
SP86	h+ *MGG-sde2-C-6HA:Nat-NT2 cdc5-9MYC:HphNT1*	This study
SP88	h+ *sde2-6HA:Nat-NT2 cwf21-9MYC:HphNT1*	This study
SP89	h+ *sde2-6HA (AAKGG):Nat-NT2 cwf21-9MYC:HphNT1*	This study
SP90	h+ *sde2-6HA (K85M):Nat-NT2 cay1-9MYC:HphNT1*	This study
SP93	h+ *cdc5-9MYC:HphNT1 Δsde2:: Nat-NT2*	This study
SP97	h- *sde2-6HA:Nat-NT2 cay1-9MYC:HphNT1*	This study
SP107	h- *sde2-6HA (K85R):Nat-NT2 cay1-9MYC:HphNT1*	This study
SP108	h- *sde2-6HA (K85T):Nat-NT2 cay1-9MYC:HphNT1*	This study

Table 2.2 : The list of plasmid clones used in this study

Name	Description	Reference
pREP81x-3MYC–sde2–3FLAG	*S. pombe sde2* in pREP81x with 3MYC tag at the N-terminus and 3FLAG tag at the C-terminus	This study
pREP81x-3MYC–sde2(AAKGG)–3FLAG	G83A, G84A mutation of *pREP81x-3MYC-sde2-3FLAG*	This study
pREP81x-3MYC–sde2(GGAGG)–3FLAG	K85A mutation of *pREP81x-3MYC-sde2-3FLAG*	This study
pREP81x-3MYC–sde2(GGKAA)–3FLAG	G86A, G87A mutation of *pREP81x-3MYC-sde2-3FLAG*	This study
pREP81x-3MYC–sde2(GGMGG)–3FLAG	K85M mutation of *pREP81x-3MYC-sde2-3FLAG*	This study
pREP81x-3MYC–sde2$_{UBL}$	*S. pombe sde2* (aa 1-84) in *pREP81x*	This study
pCMV26 3FLAG–HsSDE2(AAKGG)–MYC	G76A G77A mutation of *pCMV26-3FLAG-HsSDE2-MYC*	This study
pCMV26-3FLAG–HsSDE2(GGAGG)–MYC	K78A mutation of *pCMV26-3FLAG-HsSDE2-MYC*	This study
pREP81x M–sde2-C–3FLAG	*S. pombe sde2* C-term (aa 86-263) starting with methionine in *pREP81x*	This study
pREP81x-3MYC– HsSDE2–3FLAG	*H. sapiens SDE2* cDNA in *pREP81x* with a 3MYC tag at the N-terminus and 3FLAG tag at the C-terminus	This study
pREP81x-Sphub1	*S. pombe hub1* cDNA in *pREP81x*	This study
pREP81x-Sphub1(I42S)	I42S mutation of *pREP81x-Sphub1*	This study
pREP81x-EGFP–sde2–3FLAG	EGFP fusion of *S. pombe sde2*-3FLAG in *pREP81x*	This study
p426ADH-3MYC–Spsde2–3FLAG	*S. pombe sde2* in *S. cerevisiae* expression vector *p426ADH*	This study
p424ADH-Scubp15	*S. cerevisiae ubp15* in *S. cerevisiae* expression vector *p424ADH*	This study

p424ADH-Spubp5	*S. pombe ubp5* cDNA in *S. cerevisiae* expression vector *p424ADH*	This study
p424ADH-Spubp5(C222S)	C222S mutation *in p424ADH-Spubp5*	This study
p424ADH-Spubp15	*S. pombe ubp15* cDNA in *S. cerevisiae* expression vector *p424ADH*	This study
p424ADH Spubp15(C239S)	C239S mutation in *p424ADH-Spubp15*	This study
pREP81x-3MYC– Spubp5	*S. pombe ubp5* cDNA in *S.pombe* expression vector *pREP81x*	This study
pREP81x-3MYC– Spubp15	*S. pombe ubp15* cDNA in *S.pombe* expression vector *pREP81x*	This study
pREP81x-3MYC–HsUSP7	*Hs USP7* cDNA in *S.pombe* expression vector Prep81x	This study
pREP81x-3MYC–sde2(GAKGG)–3FLAG	G84A mutation of *pREP81x-3MYC-sde2-3FLAG*	This study
pREP81x-3MYC–sde2(GVKGG)–3FLAG	G84V mutation of *pREP81x-3MYC-sde2-3FLAG*	This study
pREP81x-sde2$_{UBL}$–GFP	*sde2*$_{UBL}$ with EGFP in *pREP81x*	This study
pREP81x-sde2$_{UBL}$–GFP–NLS	*sde2*$_{UBL}$ with EGFP and SV40 NLS in *pREP81x*	This study
pREP81x-sde2$_{UBL}$–GFP–NES	*sde2*$_{UBL}$ with EGFP and Mex67 NES (aa 434-509) in *pREP81x*	This study
pPROEX-6HIS–sde2	*S. pombe sde2* in *pPROEX* with 6HIS tag at the N-terminus	This study
pPROEX-6HIS–ubi-sde2-c	*S. pombe* ubiquitin fused to sde2-C in *pPROEX* with 6HIS tag at the N-terminus	This study
pPROEX-6HIS–c1orf55	*Hs sde2* (c1orf55) in *pPROEX* with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp5	*S. pombe ubp5* in *pCDFduet* with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp5(C222S)	C222S mutation of *pCDFduet-6HIS-ubp5*	This study

pCDFduet-6HIS–ubp15	*S. pombe ubp15* in pCDFduet with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp15(C239S)	C239S mutation of *pCDFduet-6HIS-ubp15*	This study
pCDFduet-6HIS–ubp15	*Hs USP7* in *pCDFduet* with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–7USP-15UBL	*Hs USP7(1-537 aa)−Sp ubp15(562-1129 aa)* fusion with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp15ΔUBL	*S. pombe ubp15(1-561 aa)−60 aa nonspecific extension* in *pCDFduet* with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–U15U7	*Spubp15(1-379 aa)−HsUSP7(364-1102 aa)* fusion with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–U7U15	*HsUSP7(1-363 aa)−Spubp15(380-1129 aa)* fusion with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp15Δ10aa (518-530)	*S. pombe ubp15* with 10 amino acid *(518-530)* deletion in USP domain cloned in *pCDFduet* with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp5ΔMATH	*S. pombe ubp5* with MATH domain deleted from *(2-187 aa)* and cloned in *pCDFduet* with 6HIS tag at the N-terminus	This study
pCDFduet-6HIS–ubp5ΔUBL(529-1108)	*S. pombe ubp5* with deleted UBL domains *(555-1108 aa)* and deleted USP domain region *(529-544 aa)* cloned in *pCDFduet* with 6HIS tag at the N-terminus	This study
pENO-EGFP–sde2 (GVKGG)–3FLAG	*EGFP fusion of S. pombe sde2-3FLAG* under *eno1* promoter	This study
pGEX 5X1 GST-sde2-C	Sde2-C in *pGEX 5X* vector with GST tag at N-terminus	This study

Table 2.3- List of primers used in RT PCR and qPCR assays

Number	Name	Sequence (5'-3')
SKM_PR13	*act1* F	CCCCTAGAGCTGTATTCCC
SKM_PR14	*act1* R	CCAGTGGTACGACCAGAGG
SKM_PR323	*psf3* Ex3 F	GTATCTATTCGGGACATAACCACAC
SKM_PR320	*psf3* Ex5 R	CTACGAAGTGGAATTTTGCC
SKM_PR327	*mcs2* Ex1 F	GCACTTTCTTCCGCTCTTTCC
SKM_PR326	*mcs2* Ex3 R	TTTCGGAAGCACTGTTTGACAATC
SKM_PR388	*rap1* Ex1 F	CCAAAAGCGATGGCTCGTCC
SKM_PR166	*rap1* Ex3 R	AACCGAAGCAGACTTGGAAATC
SKM_PR69	*sde2 C-term* F	AAGGGTGGTTTTGGCAGTC
SKM_PR746	*sde2 C-term* R	CTCCTGCTCATTGCGCCTTCTTGACATACG
SKM_PR1089	*ftp105* Ex1 F	ATGGGAGGCCAAGAGTCAAA
SKM_PR1095	*ftp105* Ex4 R	GAGATGTCAATACAAGCAGC

Results

3.1 Sde2 co-purifies spliceosomal proteins

To study the role of Sde2 and its associated proteins, immunoprecipitation of Sde2 complexes was done using an epitope-tagged (3X FLAG) version of Sde2 protein followed by mass-spectrometric identification of co-immunoprecipitated (Co-IP) proteins (data obtained from Shravan Kumar Mishra and Kiran Kumar Kolathur). In the analysis, the interaction between Sde2 and components of the spliceosome were observed (Fig. 3.1) as recently also reported by others [184,193].

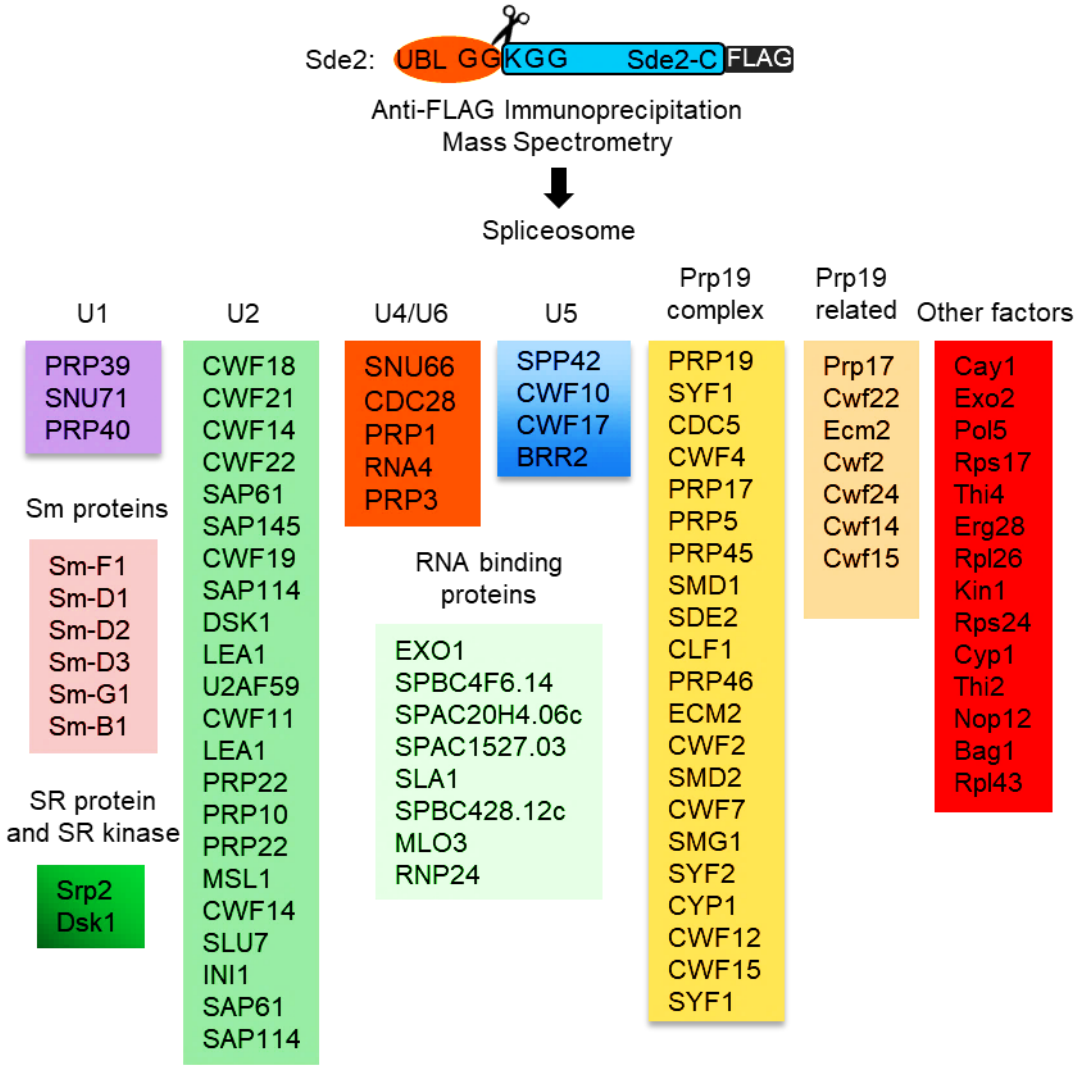


Figure 3.1 - Sde2 associates with spliceosome
Sde2 coimmunoprecipitates (CoIP) splicing factors. Lysate from *S. pombe Δsde2* strain expressing Sde2–3FLAG epitope-tagged protein was immunoprecipitated using anti-Flag antibody beads. The co-immunoprecipitated proteins were analyzed by mass spectrometry (Data is from Shravan Kumar Mishra and Kiran Kumar Kolathur).

3.2 Sde2 is processed like ubiquitin precursors

To monitor the expression of Sde2 in *S. pombe*, *sde2* gene was cloned with an epitope tag at N-terminus (3MYC) and C-terminus (3FLAG) in pREP81X vector. The plasmid had a thiamine repressible nmt promoter and Sde2 expression was induced in the absence of thiamine (EMM-leu media). To check whether Sde2 gets processed at the GGKGG motif, we sequentially mutated the residues to Alanine by site-directed mutagenesis. The mutants of *sde2* namely *AAKGG*, *GGAGG*, and *GGKAA* along with the WT were expressed in *S. pombe* (Fig. 3.2 A), and Sde2 was detected by anti-MYC and anti-FLAG immunoblotting. These experiments revealed that the full-length protein was cleaved, separating it into $Sde2_{UBL}$ and Sde2-C (lane 2 of anti-MYC and anti-FLAG immunoblots Fig. 3.2 A). Processing of Sde2 is presumed to occur at the GG^KGG sequence as alanine substitutions of the first GG residues (*AAKGG*) abolished the cleavage (lane 3 of anti-MYC and anti-FLAG immunoblots Fig. 3.2 A). By contrast, alanine mutations of the lysine (*GGAGG*) or the second GG residues (*GGKAA*) had negligible effects on the processing (lane 4 & 5 of anti-MYC and anti-FLAG immunoblots Fig. 3.2 A).

From previous reports, we know that deletion of Sde2 shows temperature-sensitive phenotype in *S. pombe*. To study the effect of Sde2 processing on its function, Sde2 GGKGG variants were transformed in *Δsde2* strain and spotted on media with and without thiamine and kept at 30°C and 37°C for 3-4 days. The ability of Sde2 variants to complement *Δsde2* growth phenotype was monitored (Fig. 3.2 B). Importantly, the alanine substitutions of the first GG residues (*AAKGG*) failed to rescue growth defects of *Δsde2* strain (Fig. 3.2 B), suggesting that cleavage of Sde2 is essential for its function. While *sde2 GGAGG* and *GGKAA* mutants were able to rescue growth defects of *Δsde2* upon overexpression, however they complemented albeit to a lesser extent than the wild type (Fig. 3.2 B).

In a parallel study (performed by my colleague Poonam Thakran) carried out on *de novo* expressions of $Sde2_{UBL}$ and Sde2-C (using a methionine start codon) for complementation of *Δsde2* strain. Here, $Sde2_{UBL}$ could not rescue growth defects in *Δsde2* strain whereas Sde2-C complemented the defects partially upon overexpression (Fig. 3.2 C). Co-expression of $Sde2_{UBL}$ and Sde2-C did not further improve growth phenotype of the deletion strain. (Fig. 3.2 C). Thus, Sde2-C is the functional unit which gets activated upon processing after $Sde2_{UBL}$. Sde2 must be synthesized in the precursor form that should get processed for its function.

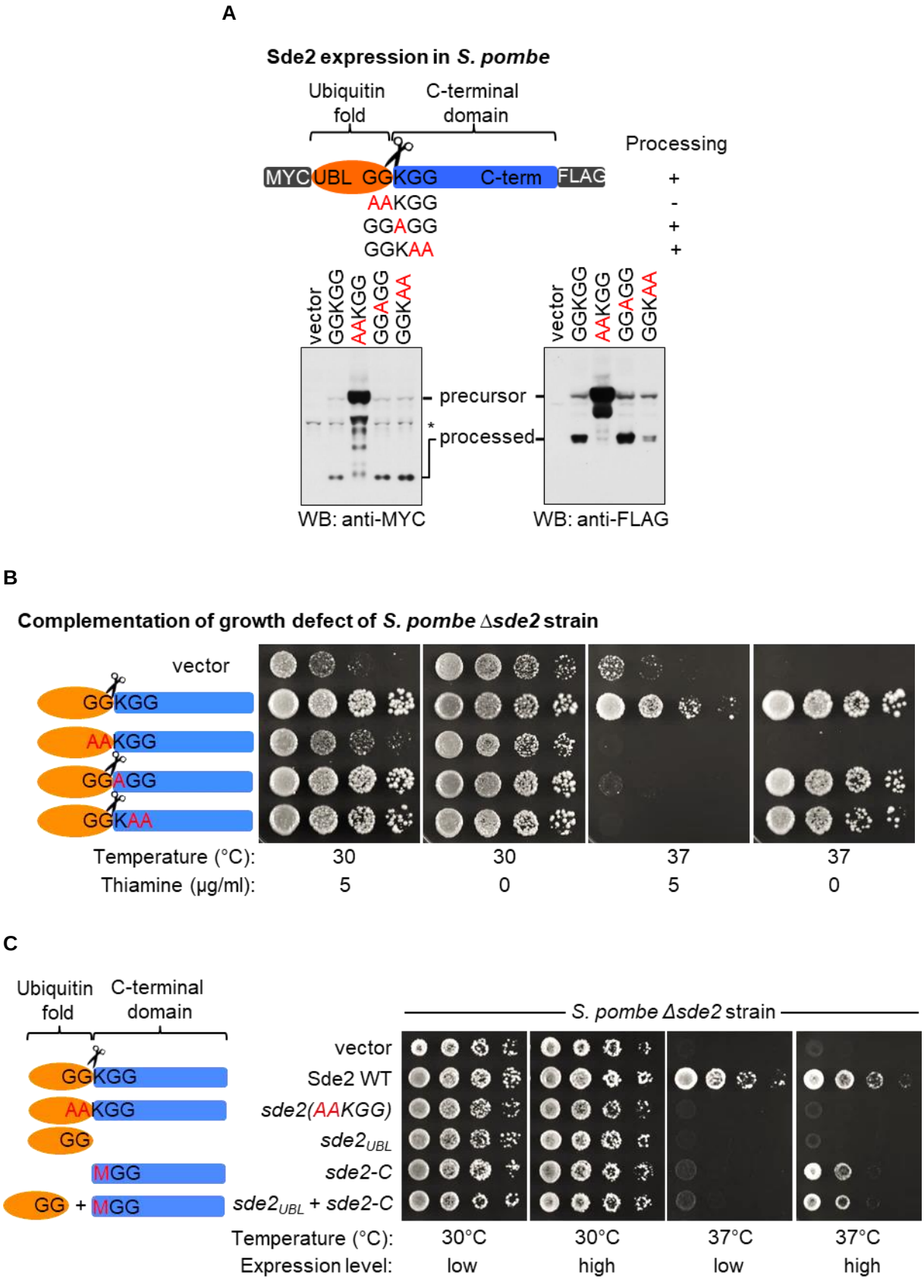


Figure 3.2 – Processing of Sde2 is required for its function. (Description of the figure is on the next page)

Figure 3.2 – Processing of Sde2 is required for its function.

A. Expression and processing of Sde2 protein in *S. pombe* detected by immunoblot analysis (western blot, WB) using epitope-tag specific antibodies. Red coloured residues mark changes from wild type (WT) Sde2. (∗) – indicates the non-specific band appeared in anti-MYC WB.
B. Complementation of *S. pombe Δsde2* by GGKGG mutants of Sde2. Constructs are as in A. Five-fold serial dilution spotting was done on indicated agar plates. Plates were incubated at 30°C and 37°C until growth appeared.
C. Complementation of *S. pombe Δsde2* by Sde2 domains. The experiment is as in C. Expression constructs encoding Sde2 WT, the processing defective mutant *sde2 (AAKGG),* $Sde2_{UBL}$, Sde2-C, and $Sde2_{UBL}$ and Sde2-C together were used.

3.3 Sde2 is processed in the nucleus

To check the subcellular localization of Sde2-C, we chromosomally tagged Sde2 with GFP at the C-terminus (Fig.3.3 B), by using this strain we wanted to see the localization of Sde2-C. We made additional clones where GFP was fused at the N-terminus of WT Sde2 (Fig.3.3 A) and *sde2 GVKGG*, the processing defective mutant (Fig.3.3 C). GFP–Sde2 and GFP–*sde2 GVKGG* chimeras were kept under thiamine repressible promoter and ENO promoter respectively. These plasmid-borne clones were used for localization studies of $Sde2_{UBL}$ and the precursor. After processing, $Sde2_{UBL}$ was diffusely localized in *S. pombe* (Fig.3.3 A) whereas Sde2-C and the precursor were predominantly nuclear (Fig.3.3 B & C). We looked for an internal nuclear localization signal in Sde2 using NLStradamus online tool [194], that predicted a sequence in Sde2-C from 135-KKPAETRAKKEAKKQK-150 amino acids to be a probable NLS sequence at a 0.6 threshold. After synthesis, Sde2 may directly diffuse to the nucleus or because of NLS like sequence in Sde2-C, is targeted to the nuclear compartment. From our localization studies of Sde2 and also evident from previously reported literature [195], Sde2 is localized predominantly in the nucleus.

To monitor the sub-cellular location of Sde2 processing, we generated $Sde2_{UBLGG}$–GFP chimera, with a nuclear localization signal (NLS) sequence of SV40 large T antigen (PKKKRKV) and two nuclear export signal (NES) sequences of Mex67 gene of *S. pombe* at the C-terminus of GFP (Fig.3.3 D). Mex67 protein functions in the export of polyA mRNA and large ribosomal subunits from the nucleus to the cytoplasm. The clones were expressed under nmt81x promoter for controlled expression of the chimeras in *S. pombe*. We found that $Sde2_{UBLGG}$–GFP, and $Sde2_{UBLGG}$–GFP–NLS chimeras were processed efficiently whereas processing of $Sde2_{UBLGG}$–GFP–NES was significantly affected (Fig.3.3 D). The results suggest that processing of $Sde2_{UBLGG}$–GFP–NES chimera is hampered since the NES sequence keeps the majority of the protein in the cytoplasm. $Sde2_{UBLGG}$–GFP–NLS chimera

is efficiently processed when localized in the nucleus. The GFP alone is known to localize all over the cell when expressed in *S. pombe* [196] and is not restricted to any specific cell compartment.

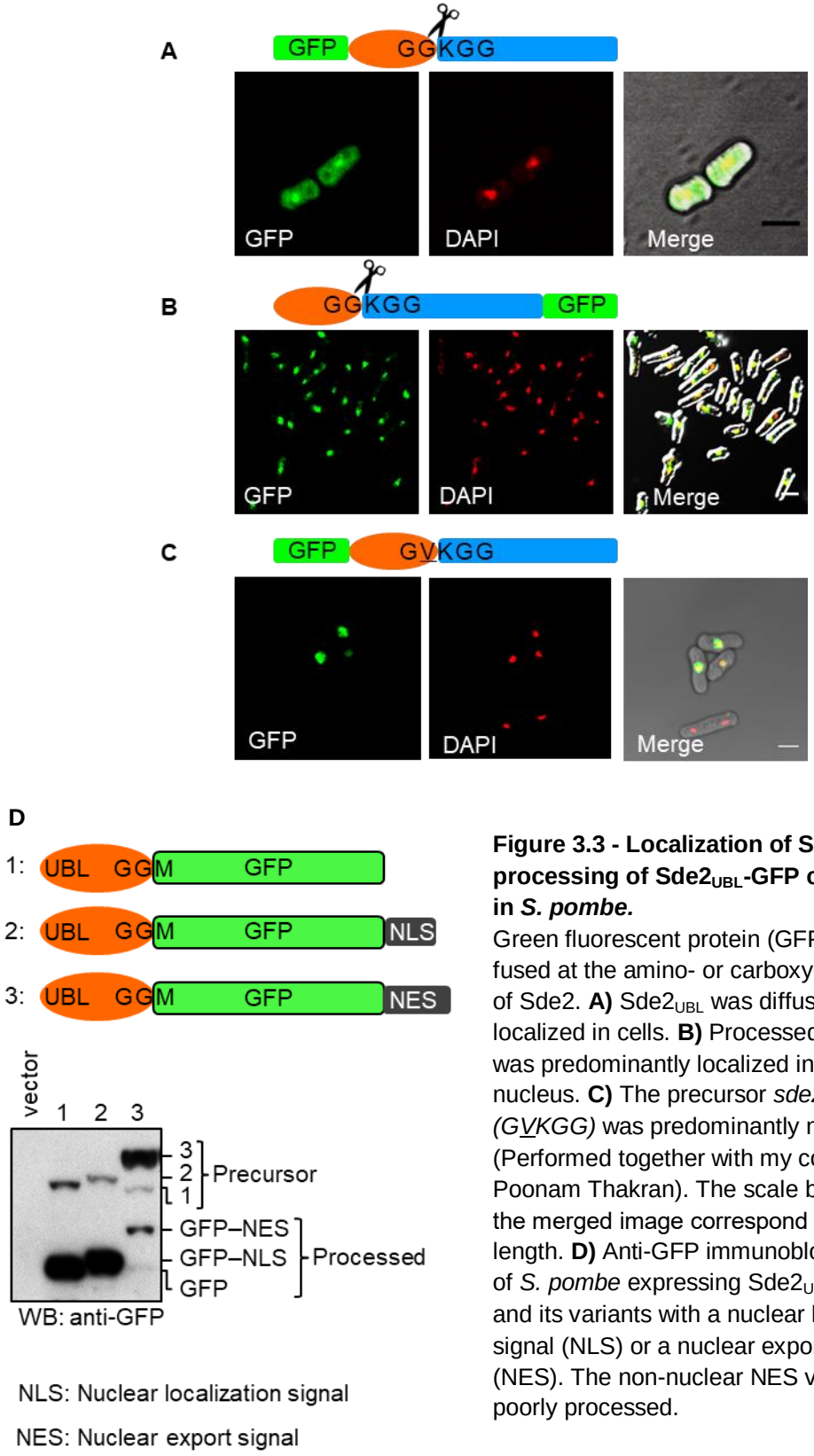


Figure 3.3 - Localization of Sde2 and processing of Sde2$_{UBL}$-GFP chimeras in *S. pombe*.
Green fluorescent protein (GFP) was fused at the amino- or carboxy-terminus of Sde2. **A)** Sde2$_{UBL}$ was diffusely localized in cells. **B)** Processed Sde2-C was predominantly localized in the nucleus. **C)** The precursor *sde2 (GVKGG)* was predominantly nuclear. (Performed together with my colleague Poonam Thakran). The scale bar size in the merged image correspond to 5 µm length. **D)** Anti-GFP immunoblot analysis of *S. pombe* expressing Sde2$_{UBL}$–GFP, and its variants with a nuclear localization signal (NLS) or a nuclear export signal (NES). The non-nuclear NES version was poorly processed.

Therefore, $Sde2_{UBL}$–GFP might be localized all over the cell, and the nuclear fraction of $Sde2_{UBL}$–GFP pool has got processed. There is a subtle difference between processing efficiency of $Sde2_{UBL}$–GFP and $Sde2_{UBL}$–GFP–NLS since major fraction of the later is in the nucleus.

In conclusion, the results of Sde2 being predominantly localized in the nucleus, poor processing of $Sde2_{UBL}$–GFP–NES chimera in the cytoplasm in contrast to efficient processing of $Sde2_{UBL}$–GFP–NLS chimera in the nucleus indicates that the processing of Sde2 takes place in the nucleus.

3.4 Sde2 precursor accumulates in Δ*ubp5* Δ*ubp15* strain

To find out the processing enzyme of Sde2, we searched for Sde2 interactors from the proteomics data for an upregulated protein with protease activity. We also performed yeast two-hybrid screen using *S. pombe* genes two-hybrid library with s*de2 (AAKGG)* processing defective mutant. However, we could not find a protease that interacted directly with Sde2. Therefore, we decided to monitor the accumulation of Sde2 precursor in selected mutants of *S. pombe* proteases by immunoblot assays. The candidate protease deletion strains were selected by similarity to growth phenotype of Δ*sde2* strain and by presence or absence of protease homologs in *S. cerevisiae*. A criterion was used that, if any protease deletion strain shows Sde2 precursor accumulation with a relative reduction in processed form, will be selected as candidate protease for further studies. The deletion strains were revived from Bioneer haploid gene deletion library of *S. pombe*. The processing defective *sde2 (AAKGG)* mutant was taken as a control to monitor precursor accumulation. The proteasomal mutants *mts2-1* and *mts3-1* were also taken to check the possibility of Sde2 processing at the proteasome. Mts2 and Mts3 are part of 19S proteasome subunit complex, and its mutants show defects in proteasomal degradation at higher temperature [197]. For *mts2-1* and *mts3-1* strains, Sde2 transformed cells were grown till the logarithmic phase and 1 OD cells were harvested at 0 hr time point. The cells were further grown at higher temperature (37°C) for giving heat shock and again 1 OD cells were harvested after 1 and 3 hrs.

The precursor accumulated in a strain lacking *ubp15* gene, a ubiquitin-specific protease (USP) domain-containing DUB (Fig.3.4 A) Upon repetition of the experiment with Δ*ubp15* and WT strains, we observed residual processing of Sde2-C in *Δubp15* strain, suggesting the involvement of additional enzymes for the processing of Sde2. It was previously reported that Ubp15 genetically interacts with its paralogue Ubp5 and the double

mutant shows synthetic sickness [198]. However, in the screen, we did not see any accumulation of the precursor in Δ*ubp5* strain.

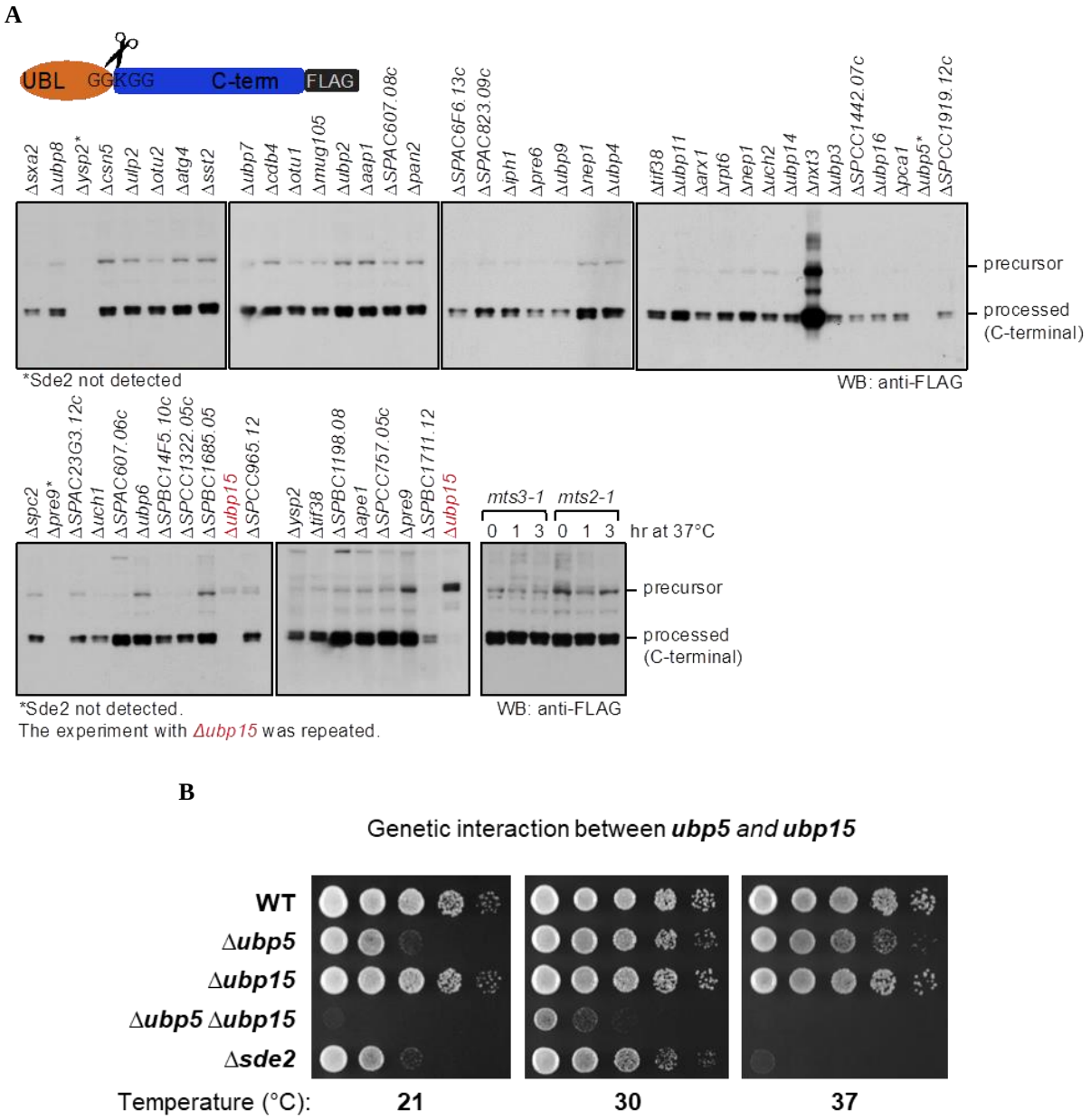


Figure 3.4 - Screening for Sde2 protease.
A) The strains are haploid deletion mutants of proteases not essential for viability. The temperature sensitive proteasome mutants *mts2-1* and *mts3-1* were also tested for processing of Sde2. WT *sde2–3FLAG* construct was expressed in all the strains. (∗) The experiment was repeated for strains marked with asterisk since Sde2 was not detected in the first attempt. Cell lysates were analyzed by western blot analysis using anti-FLAG antibody. Accumulation of Sde2 precursor in *Δubp15* strain indicates that this enzyme is required for processing of the protein in *S. pombe*. Ubp15 is a ubiquitin-specific protease (USP) domain containing deubiquitinating (DUB) enzyme [52]. **B)** Genetic interaction between *Ubp5* and *Ubp15. Δsde2* strain does not phenocopy the deletion strains of the proteases *Δubp5*, *Δubp15 or Δubp5 Δubp15*. Spot assay was performed on YES media and incubated at temperature conditions mentioned at the bottom of each spot assay.

So, we deleted both *ubp5* and *ubp15* genes and observed growth phenotypes of single and double deletion strains at unfavourable temperature conditions (Fig.3.4 B). *Δubp5 Δubp15* double mutant showed severe growth phenotype at 21°C, 30°C and 37°C (Fig.3.4 B), while *Δubp5* strain showed poor growth at 21°C and 37°C.

To establish the connection of genetic interaction of Ubp5 and Ubp15 with processing defect in Sde2, we expressed Sde2 WT and *AAKGG* mutants in *S. pombe* WT, *Δubp5, Δubp15*, and *Δubp5 Δubp15* strains (Fig.3.4.1A). The processing of Sde2 in these strains was monitored by anti-FLAG immunoblotting. We observed that Sde2 processing was abolished entirely in *Δubp5 Δubp15* double mutant whereas it was partially processed in *Δubp15*. Sde2 was optimally processed in WT and *Δubp5* mutant strains. For detection of endogenous Sde2, we used anti-Sde2 antibody conjugated beads and performed immunoprecipitation experiment in *S. pombe* WT, *Δsde2, Δubp5, Δubp15, and Δubp5 Δubp15* strains.

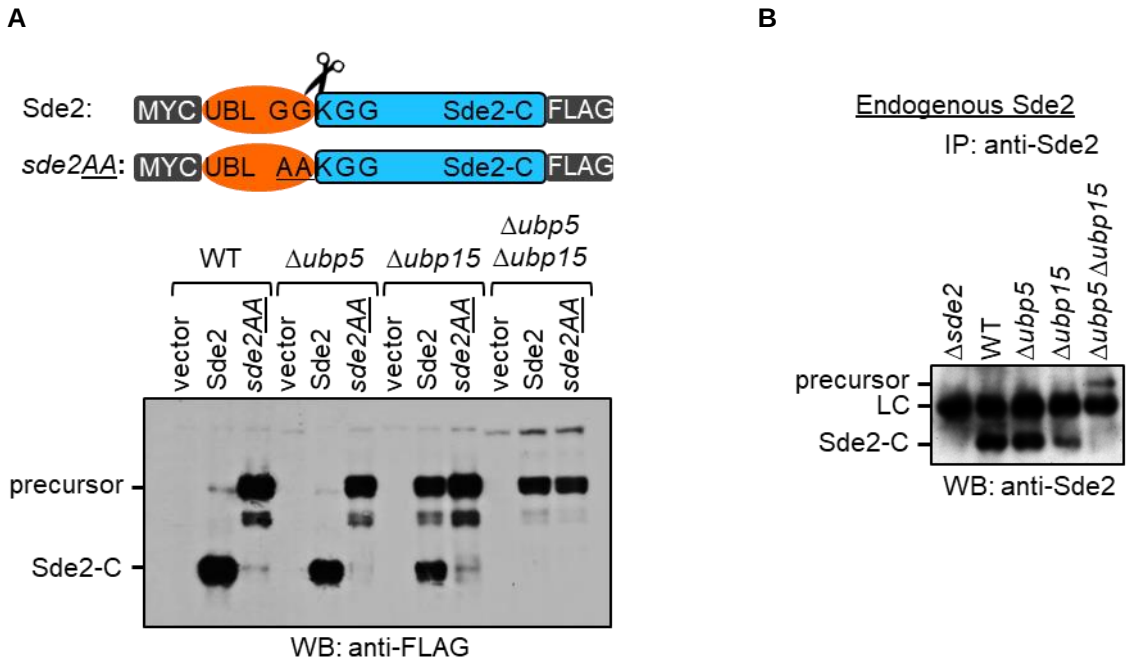


Figure 3.4.1 - Sde2 precursor accumulates in Δ*Ubp5* Δ*Ubp15* strain

A. Sde2 processing in *S. pombe*. Constructs from Fig 1B were expressed in *S. pombe* WT, Δ*ubp5*, Δ*ubp15,* and Δ*ubp5* Δ*ubp15* strains. The protein expression was analyzed by anti-FLAG immunoblotting.

B. Endogenous *S. pombe* Sde2-C and the full-length Sde2 precursor in Δ*ubp5* Δ*ubp15* strain. Sde2 protein from indicated strains was immunoprecipitated using a polyclonal antibody against recombinant Sde2-C, followed by western blot analysis using the same antibody. LC – indicates the Light Chain of anti-Sde2 antibody in the figure.

We could detect endogenous Sde2-C in WT, *Δubp5*, *Δubp15* strains whereas, accumulation of full-length Sde2 precursor was observed in *Δubp5 Δubp15* double mutant (lane 5 in Fig.3.4.1B). Ubp5 and Ubp15 are generally involved in the processing of ubiquitin and ubiquitin conjugates [52], and clearly Sde2 is not the only substrate of Ubp5 and Ubp15. Accordingly, the *Δubp5 Δubp15* double deletion resulted in elevated growth defects, i.e., due to Sde2 dysfunction as well as accumulation of ubiquitin-conjugated substrates targeted by these DUBs. Thus, *Δubp5 Δubp15* deletion strain showed more severe growth defects than *Δsde2* strain.

3.5 Ubiquitin-specific proteases Ubp5 and Ubp15 process *S. pombe* Sde2

To further test whether Ubp5 and Ubp15 acted directly as the proteases, we examined the processing of *Sp*Sde2 by these DUBs in the recombinant system of *Escherichia coli*. Efficient processing of Sde2 was readily observed in *E. coli* upon co-expression of Ubp5 (lane 4 in Fig. 3.5) and Ubp15 (lane 6 in Fig. 3.5). USP family DUBs belong to cysteine proteases which use active site cysteine for nucleophilic attack on isopeptide bond while processing ubiquitin conjugates.

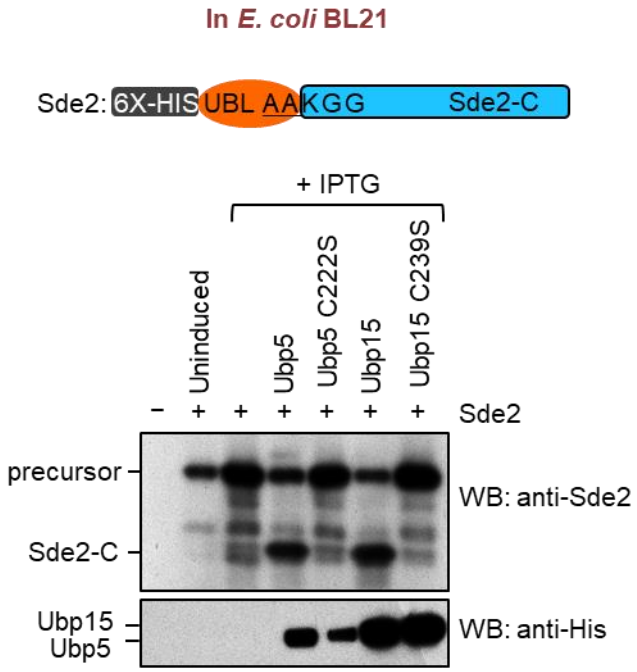


Figure 3.5 - Processing of Sde2 by Ubp5 and Ubp15 in a recombinant system
Expression constructs harbouring indicated cDNAs were co-transformed in *E. coli* BL21 (DE3) strain. Following protein expression, total cell lysates were processed by immunoblotting using anti-Sde2-C antibody. The proteases expression was checked by the anti-His antibody.

To confirm that Ubp5 and Ubp15 cleave Sde2 through their active site cysteine residues, we mutated cysteine to serine in both the enzymes. The catalytically inactive cysteine mutants of Ubp5 C222S and Ubp5 C239S were unable to cleave Sde2 (land 5 and 7

in Fig. 3.5), indicating that these DUBs directly process Sde2 via the active site present in USP domain.

Although *S. cerevisiae* lacks Sde2, it contains an Ubp15 ortholog known to be active on ubiquitin. To check whether *Sc*Ubp15 has retained the ability to process *Sp*Sde2 we co-expressed them in *S. cerevisiae*. Interestingly, *Sp*Sde2 was poorly processed in *S. cerevisiae* by *Sc*Ubp15 which was apparent upon overexpression of *Sc*Ubp15 (lane 3 of long exposure immunoblot in Fig. 3.5.1). The processing of *Sp*Sde2 was not visualized with endogenous *Sc*Ubp15 by immunoblot assay (lane 2 in Fig. 3.5.1).

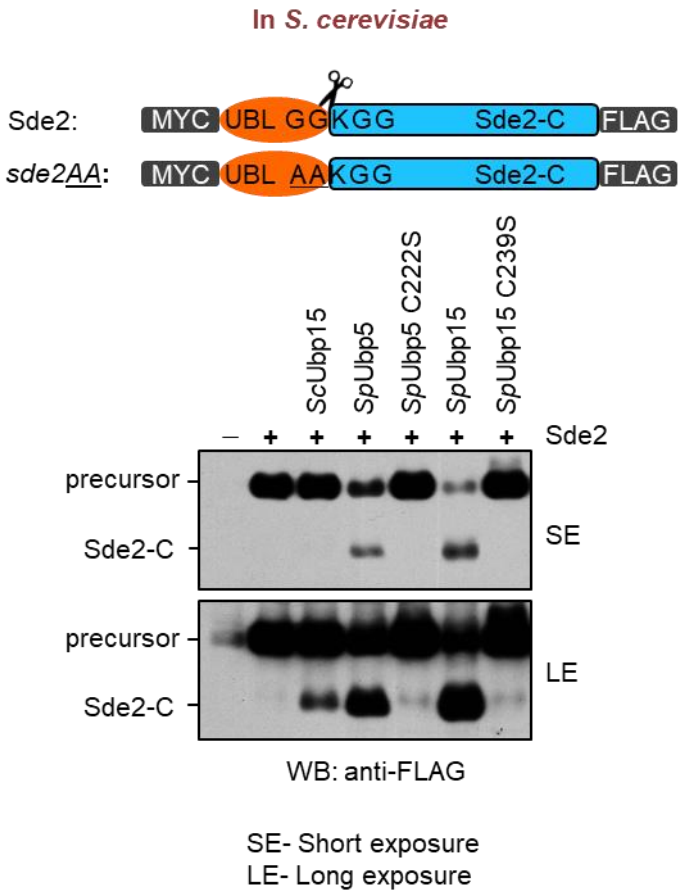


Figure 3.5.1 – Processing of *S. pombe* Sde2 in the budding yeast *S. cerevisiae*
Yeast expression constructs harboring *S. pombe sde2* full-length cDNA with 3FLAG epitope tag at its C-terminus under a constitutive *ADH* promoter is used. For expression of *S. cerevisiae UBP15*, *S. pombe ubp5* and *ubp15*, constructs cloned in the p424*ADH* vector were used. The active site cysteine to serine mutants of *S. pombe* enzymes used are *Sp*Ubp5 C222S and *Sp*Ubp15 C239S. Processing of Sde2 was analyzed by anti-FLAG western blot.

The poor processing of Sde2 by *Sc*Ubp15 is presumably because *Sp*Ubp15 and *Sc*Ubp15 share only 44% identity and secondly due to altered localization of *Sp*Sde2 in *S. cerevisiae*. It is reported that *Sc*Ubp15 localizes in cytosol and peroxisomes [199]. Upon co-

expression of *Sc*Ubp15 with *Sp*Sde2 in *S. cerevisiae*, how likely they co-localize for processing needs further studies. It will become clear upon co-expression in recombinant system like *E. coli* that how efficiently *Sc*Ubp15 is able to process *Sp*Sde2. Importantly, increased efficiency of *Sp*Sde2 processing was observed upon co-expression of either *Sp*Ubp5 or *Sp*Ubp15 in *S. cerevisiae*, highlighting their role in the processing of *Sp*Sde2 (lane 4 and 6 in Fig.3.5.1).

To confirm whether Sde2 processing ability of Ubp5 and Ubp15 is a specific or general property of USP domain-containing proteases, we cloned another DUB Ubp16 from *S. pombe*. It is a 457 amino acid long USP domain-containing enzyme and has 27% identity to Ubp15's USP domain. Co-expression of Ubp15 and Ubp16 with Sde2 and Ub-Sde2-C in *E. coli* revealed that Ubp15 processed both Sde2 and ubiquitin whereas; Ubp16 could only process ubiquitin fused substrate (Fig.3.5.2).

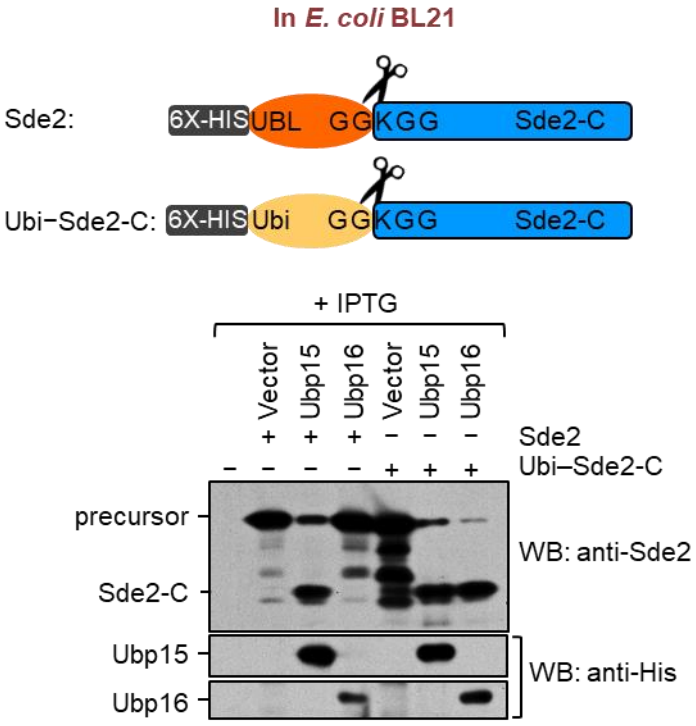


Figure 3.5.2 - A USP domain-containing DUB Ubp16 does not process Sde2.

Another USP domain-containing DUB Ubp16 does not process Sde2. Experiment similar to (Fig.3.6), Processing of ubiquitin–Sde2-C fusion by these proteases was used as a control. Anti-HIS immunoblotting was used to monitor expression of the proteases.

Presumably, other DUBs might have lost the ability to process Sde2, or $Sde2_{UBL}$ evolved to become specific only for Ubp5 and Ubp15. Thus, the activity of these proteases on Sde2 is highly specific, as another USP domain-containing DUB Ubp16, did not process Sde2.

3.6 *Δubp5 Δubp15* strain shows *Δsde2* like splicing defects

To elucidate the role of Sde2 in pre-mRNA splicing, we analyzed splicing defects in *Δsde2* strain using *S. pombe* genome-wide splicing-sensitive microarrays [192]. The experiment was done by Shravan Kumar Mishra in collaboration with Jeffrey A. Pleiss at Cornell University. The microarrays were designed to assess the splicing defects in nearly all intron-containing genes in *S. pombe*. The arrays have oligonucleotide probes specific for exons, introns and exon-exon junctions for nearly all intron-containing genes in *S. pombe*. The probes for exons quantify the signal of total transcripts of a gene, intronic probes measure the level of unspliced transcripts and the exon-exon junction probes are for quantification of spliced product/mRNA of the given genes.

Since *S. pombe Δsde2* strain showed temperature-sensitive phenotype or poor growth at 37°C, we used this strain to study Sde2's role in pre-mRNA splicing. The total RNA was isolated from *S. pombe* wild-type and *Δsde2* strains grown at 30°C (optimal temperature for growth) or after shifting to 37°C (unfavourable temperature for growth) for 15 minutes to detect early splicing defects. The RNA was converted into cDNA using random nine-mer primers and reverse transcriptase. The cDNAs from both wild-type and *Δsde2* strain were labeled with Cy3, Cy5, and dye-swapped and hybridized on splicing-sensitive microarrays. After hybridization relative abundance of all transcripts were quantified and compared in both strains.

Splicing of the most introns was unaffected in *Δsde2* strain under either growth condition, as evidenced by the lack of apparent changes in pre-mRNA levels. Interestingly, however, splicing defects were observed for nearly 30 genes, where pre-mRNA levels were increased by at least two-fold and were often accompanied by decrease in the levels of mature mRNA. The list of genes where splicing defects were observed for specific introns is shown in the appendix section (Table 5.1). Splicing defects in *Δsde2* strain were confirmed by RT-PCR assays, followed by agarose gel electrophoresis to visualize intron-containing transcripts. Key targets of Sde2 include multi-intronic genes *rap1* (encodes a telomere binding protein with function in telomeric silencing), *psf3* (encodes a GINS complex subunit with function in DNA replication), and *mcs2* (encodes a TFIIH complex cyclin with function

in transcription). Sde2 functions as an intron-specific splicing regulator is shown by my colleagues Poonam Thankran, Shravan Kumar Mishra and Jeffrey A. Pleiss [200].

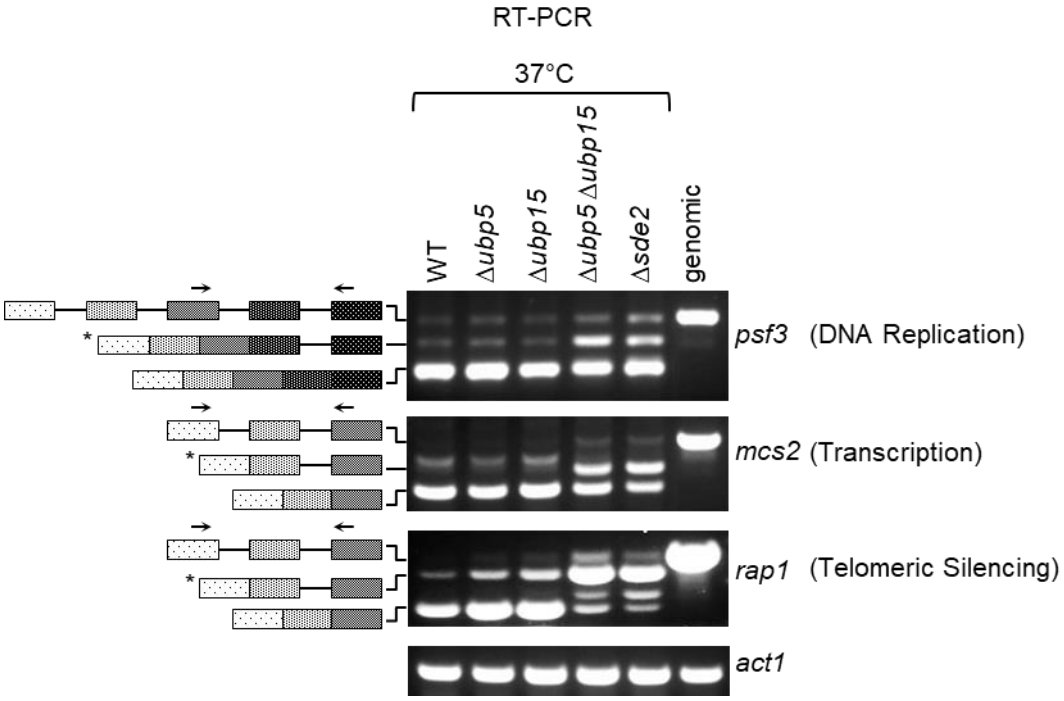


Figure 3.6 - *Δubp5 Δubp15* strain shows *Δsde2* like splicing defects
RT-PCR assays to monitor the accumulation of intron-containing transcripts in *Δubp5 Δubp15* mutant defective in Sde2 processing. *Δubp5 Δubp15* strain showed *Δsde2*-like intron-specific splicing defect 37°C. Black arrows marks primers used for RT-PCR assay. Asterisks mark intron-containing transcripts accumulated in *Δsde2* strain.

To confirm that the processing is essential for Sde2's splicing function; WT, Δ*ubp5,* Δ*ubp15,* Δ*sde2,* and Δ*ubp5* Δ*ubp15* double mutants were grown till early logarithmic phase (0.5 OD_{600}) and subjected to heat shock at 37°C, 15 min, followed by cell harvesting and RNA isolation. In correlation with Δ*sde2,* transcripts having retained only one of the intron (for *rap1* intron-2, *psf3* intron-4, and *mcs2* intron-2), accumulated in Δ*ubp5* Δ*ubp15* strain whereas individual protease deletion strains showed splicing profile like WT (Fig.3.6). Thus the DUBs Ubp5 and Ubp15 regulate the splicing by processing Sde2 precursor so that activated Sde2-C becomes available for splicing function.

3.7 Processing of $Sde2_{UBL}$ promote association of Sde2-C with the spliceosome

To better understand the cellular role of Sde2, mass spectrometry was used to identify proteins that co-immunoprecipitated (CoIP) with an epitope-tagged version of Sde2 (Fig.3.1). Here, we confirmed the association between Sde2 and components of the spliceosome, as recently reported by others [184,187,200]. CoIP assays using chromosomally epitope-tagged versions of multiple Cdc5/Prp19 complex subunits, including Cdc5, Prp19, and Isy1 further validated these interactions (done by my colleague Poonam Thakran).

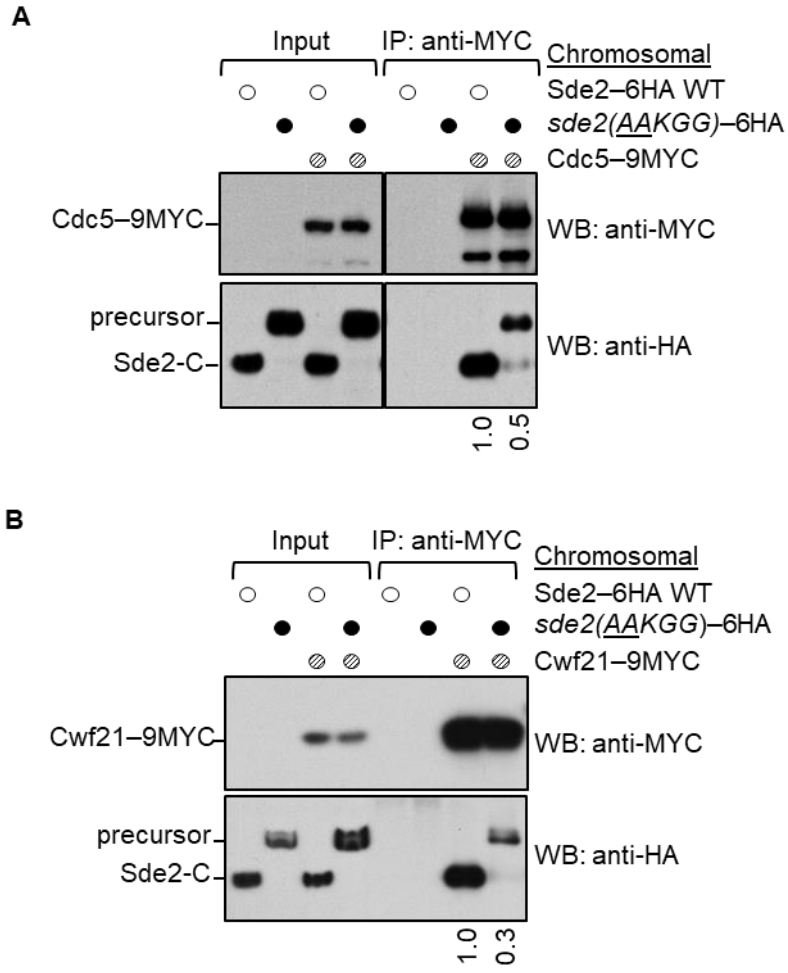


Figure 3.7 - Association of processing-defective Sde2 with the spliceosome

A. CoIP assay to monitor the association of the processing defective Sde2 with Cdc5. Experiments are performed with proteins expressed at their endogenous levels by using chromosomally tagged strains with given epitopes at the C-termini of indicated genes. IP was performed using anti-MYC agarose resins followed by immunoblot assays with anti-MYC antibody to monitor IP efficiency of Cdc5, and with anti-HA antibody to monitor CoIP of Sde2 versions. Numbers below anti-HA blot indicate the ratio of Sde2 to Cdc5 (HA/MYC) signals obtained from ImageJ quantification of immunoblot signals. Cdc5 association of the processing defective Sde2 precursor is diminished to half.

B. CoIP assay to monitor the association of the processing defective Sde2 with Cwf21. The assay is similar to (A). Cwf21 association of the processing defective Sde2 precursor is diminished to one third.

To check whether processing of Sde2 affects its association with the spliceosome, we used a technique where Sde2 WT and processing defective versions were co-IPed using spliceosomal proteins Cdc5 and Cwf21. To perform this experiment, we chromosomally inserted a sequence encoding 6HA (Hemaglutinin) epitope tag at the C-terminus of WT and processing-defective (*AAKGG*) mutant of Sde2. In each strain, we chromosomally tagged core spliceosomal factors; Cdc5 and Cwf21 with 9MYC epitope tag at the C-terminus. The coimmunoprecipitation (CoIP) of Cdc5 and Cwf21 was followed by detection of Sde2 using anti-HA immunoblotting. By contrast to the ^{Lys}Sde2-C, which associated efficiently with Cdc5 and Cwf21, the CoIP of s*de2 AAKGG* mutant was strongly diminished (Fig.3.7 A & B). The results indicate that Sde2-C's incorporation into the spliceosome is facilitated by processing of Sde2.

3.7.1 Sde2$_{UBL}$ is inhibitory for the incorporation of Sde2-C into the spliceosome

In wild type *S. pombe*, the processed form of Sde2-C is present in larger amounts than its precursor. In the processing defective mutant strains in Figures 3.7A and 3.7B, we observed a residual incorporation of Sde2 precursor in the spliceosome. Thus, we wanted to test whether this residual incorporation happened because the processed Sde2-C was not available in the mutant. We used the plasmid-borne clone of *GAKGG* mutant, whose processing is affected only partially, and generates similar levels of both precursor and processed Sde2. Additionally, the processing defective mutants *AAKGG* and *GVKGG* were used which remained largely in the precursor form and Sde2 wild-type was taken as a control. We transformed these plasmids in *S. pombe* Δ*sde2* strain where Cdc5 was chromosomally tagged with 6HA epitope tag at the C-terminus. The CoIP of Cdc5-6HA with Sde2 WT, processing defective *AAKGG*, *GVKGG*, and partially processed *GAKGG* mutants was performed to check Sde2 precursor levels incorporated into the spliceosome.

The *GAKGG* mutant produced similar levels of precursor and processed proteins (lane 3 of INPUT samples in Fig.3.7.1). Although the *AAKGG* and *GVKGG* mutants were defective in processing, fractions of the processed Sde2-C were visualized in immunoblots of immunoprecipitated samples. Herein, only the processed Sde2-C associated with Cdc5 irrespective of the mutation (IP: anti-HA samples in Fig.3.7.1). Even upon overexposure of immunoblots only processed form (Sde2-C) was visualized (IP: anti-HA samples in

Fig.3.7.1). Thus, the spliceosomal association of a small amount of Sde2 precursor is possible, but only in the absence of processed Sde2-C. Hence, $Sde2_{UBL}$ is inhibitory for the incorporation of Sde2-C into the spliceosome.

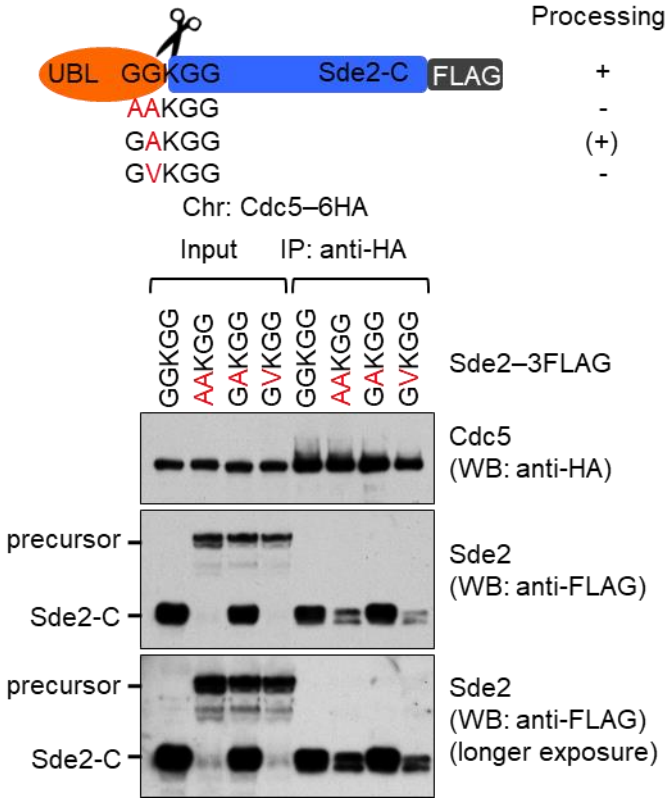


Figure 3.7.1 - $Sde2_{UBL}$ is inhibitory for the incorporation of Sde2-C into the spliceosome. Sde2 variants in the schematics were transformed into *Schizosaccharomyces pombe* Δ*sde2* with Cdc5-6HA chromosomal tag. Anti-HA IP was carried out to monitor the Co-IP of Sde2-C and precursor. Only processed Sde2-C, but not the precursor, associates with Cdc5. The plasmid clones of Sde2 GAKGG and GVKGG were made by my colleague Sumanjit Datta.

3.8 $Sde2_{UBL}$ supports optimal expression of Sde2-C protein

In complementation studies of *Δsde2* strain carried out with Sde2 GGKGG variants, i.e., the *GGAGG* mutant (Fig.3.3B) or ^{Met}Sde2-C version (Fig.3.3C) showed complementation only upon overexpression. When expressed at lower levels or in the presence of thiamine, these mutants failed to complement the growth defect of *Δsde2* strain.

The only difference between WT and mutant versions of Sde2 was the mutated N-terminal lysine residue, i.e. Alanine in *GGAGG* and Methionine in ^{Met}Sde2-C. We speculated whether these lysine mutants of Sde2-C are defective in function since Sde2-C synthesized from precursor starts with an N-terminal lysine residue, i.e., ^{Lys}Sde2-C.

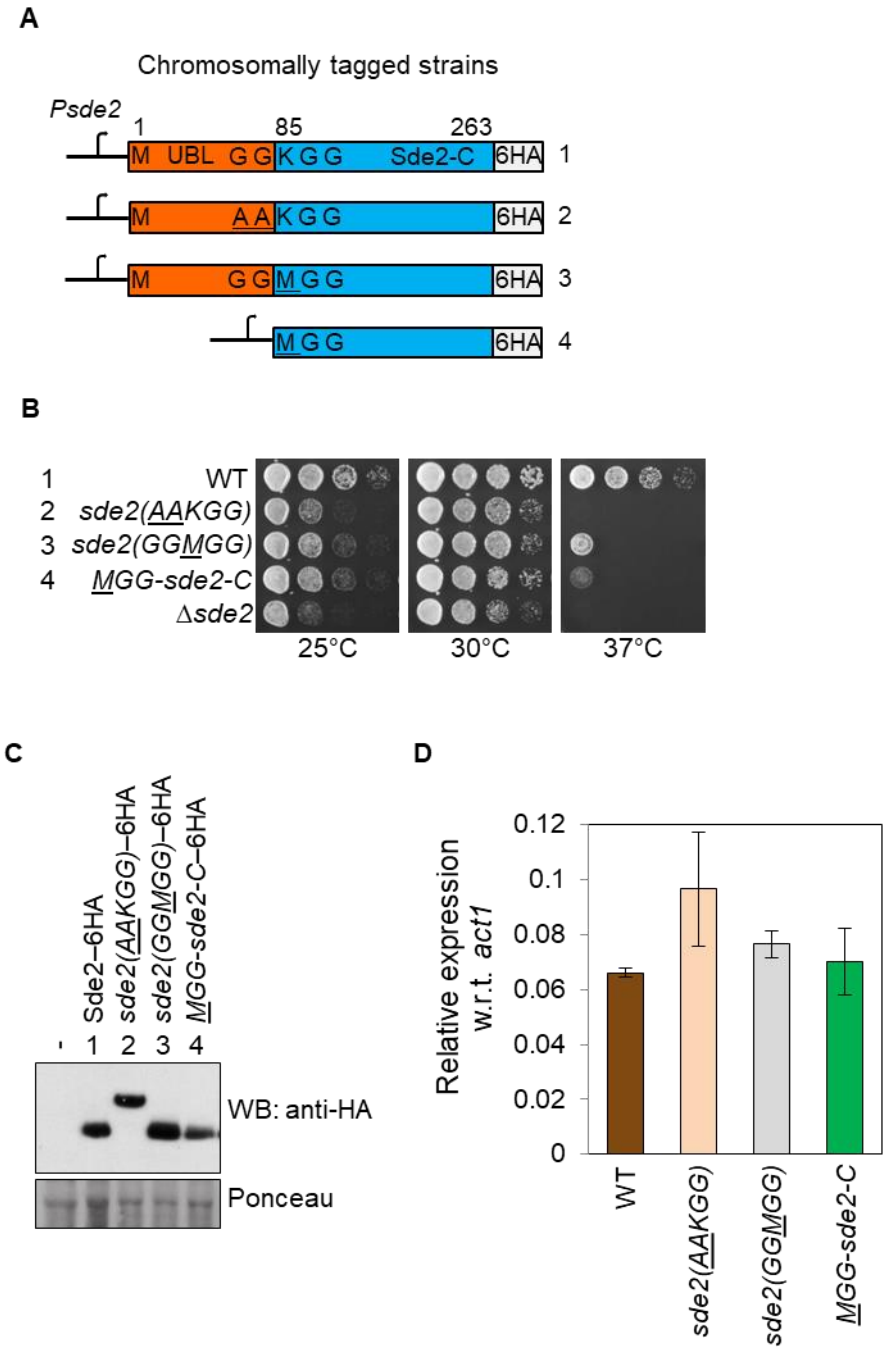


Figure 3.8 - Sde2$_{UBL}$ supports the optimal expression of Sde2-C

A. Chromosomal 6HA epitope-tagged *sde2* variants. The *sde2* promoter is common to all variants.
B. Growth phenotype of strains in (A). Strains with ^{Met}Sde2-C proteins, formed either after processing of *sde2 (GGMGG)*mutant or translated *de novo*, show Δ*sde2*-like growth defects.
C. Western blot with anti-HA antibody to detect steady-state levels of WT and mutant Sde2 proteins.
D. RT-qPCR quantification of *sde2* transcripts in strains shown in (A). RT was performed for total RNA isolated from logarithmically grown cells. *sde2* specific primers common to all *sde2* variants were used for PCR. The error bars represent the standard deviation of three PCR reactions. The *sde2* transcript level was similar in all the mutants in comparison with WT.

Also, the plasmid-borne expression may not match the endogenous levels of Sde2. Therefore, we asked whether ^{Lys}Sde2-C formed after processing at the C-terminus of Sde2$_{UBL}$ is the functional form required in the spliceosome.

To address this question, we deleted the region coding for Sde2$_{UBL}$ from chromosomal locus such that Sde2-C will start with methionine (Fig.3.8 A & B). Additionally, Sde2 WT, *AAKGG*, *GGMGG* chromosomal mutants were taken where Sde2 will be synthesized as a precursor and Sde2-C will be generated (in WT and *GGMGG*) only after cleavage at the C-terminus of Sde2$_{UBL}$. These strains were spotted along with *Δsde2* on YES media to monitor growth phenotype (Fig.3.8 B).

In growth assay, *sde2(AAKGG)* mutant showed *Δsde2* like growth phenotype at 25°C and 37°C (Fig.3.8 B, strain no. 2 and Δ*sde2*). Interestingly, both the mutant strains having ^{Met}Sde2-C showed severe growth defects at 25°C and 37°C (Fig.3.8 B, strain no. 3 and 4); however, growth of the strain having ^{Met}Sde2-C formed after processing of the *GGMGG* mutant (Fig.3.8 B, strain no. 3) was better at 37°C than with ^{Met}Sde2-C from *de novo* translation (Fig.3.8 B, strain no. 4). Further, we also checked the protein expression of all the mutants by extracting total protein from equal OD_{600} cells harvested in the logarithmic phase. The protein level was lower for ^{Met}Sde2-C from *de novo* translation (observe lane 5 of Fig.3.8 C) indicating plausible defects in transcription or translation.

To study the effect of Sde2$_{UBL}$ deletion on transcriptional synthesis, we measured the transcript levels of Sde2 WT, *AAKGG*, *GGMGG* and ^{Met}Sde2-C mutants by quantitative RT-qPCR. There was no significant difference between the transcript levels of Sde2 variants (Fig.3.8 C) indicating that transcription is not affected. Finally, we checked the half-life of these mutants to monitor protein turnover using cycloheximide shutoff assay (carried out by my colleague Sumanjit Datta) [200]. Here, ^{Lys}Sde2-C synthesized after Sde2 WT processing was a short-lived protein (Fig.3.8.1) because of the Lys at the N-terminus which acts as N-terminal degron. Such proteins with an exposed N-terminal degron are substrates of the N-end rule pathway of proteasomal degradation [201]. *GGMGG* and ^{Met}Sde2-C mutants appeared to be stable than wild-type (Fig.3.8.1) since the lysine was mutated to methionine. Thus, *de novo* translated ^{Met}Sde2-C is more stable than ^{Lys}Sde2-C indicating that Sde2$_{UBL}$ deletion is not affecting the stability of Sde2-C.

After ruling out the possibilities of the effect of Sde2$_{UBL}$ deletion on transcriptional synthesis and stability of ^{Met}Sde2-C, we could explain the reduction in *de novo* translated ^{Met}Sde2-C with two reasons. First, the translation of ^{Met}Sde2-C may get affected due to deletion of the Sde2$_{UBL}$ coding region proximal to the promoter of Sde2. Second, Sde2$_{UBL}$

might be working as a co-translational stabilizer for Sde2-C and supporting its optimal expression. As reported earlier for ubiquitin-ribosomal fusions, the ubiquitin has a co-transcriptional chaperonic function for the synthesis of the ribosomal fusion protein. We also replaced $Sde2_{UBL}$ with ubiquitin at the chromosomal locus of Sde2 and this strain did not show any growth defects (performed by my colleague Karan Chaudhary). The *in vivo* scenario also supports that ^{Lys}Sde2-C is a short-lived protein being a substrate of N-end rule pathway of protein degradation [200]. Therefore, it is likely that $Sde2_{UBL}$ functions in a similar way like ubiquitin and helps in the optimal expression of Sde2-C [3,6].

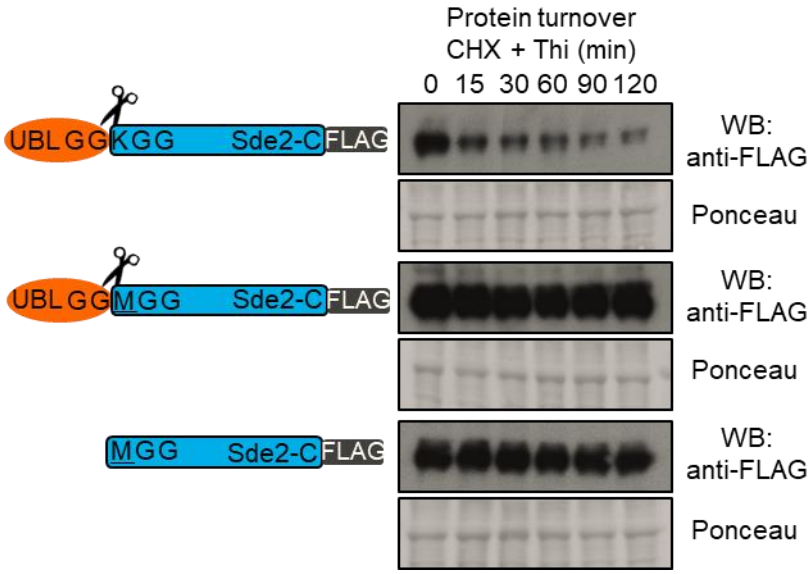


Figure 3.8.1 – Turnover of Sde2 WT, *GGMGG*, and *Met-sde2-C* mutants
Constructs used to monitor steady-state level and turnover of the proteins. Plasmids contain 3FLAG epitope tags at the C-termini of Sde2 under nmt81 promoter. Protein turnover assays - The Sde2 variants with C-terminal 3FLAG tag were expressed in *Δsde2* strain from a plasmid under *nmt81* promoter. In Sde2 WT, after processing Sde2-C starts with lysine, in *GGMGG* mutant with methionine and *Met–sde2-C* is *de novo* translated. Total proteins from 1.0 OD_{600} cells for the given time points were run on SDS-PAGE followed by anti-FLAG western blotting (This experiment was carried out by my colleague Sumanjit Datta) [200].

We also checked the spliceosomal incorporation of these mutants (figure 3.8A), where we used Cdc5-9MYC chromosomally tagged strain and generated Sde2 WT, *GGMGG* and ^{Met}Sde2-C mutants with 6HA epitope tag at the C-terminus. Cdc5 co-IP was carried out with anti-MYC beads and levels of Sde2 variants were monitored by anti-HA immunoblot assay. In Cdc5 CoIP of free ^{Met}Sde2-C, formed either after processing of the *GGMGG* precursor or from *de novo* translation, remained largely unaffected (Fig.3.8.2).

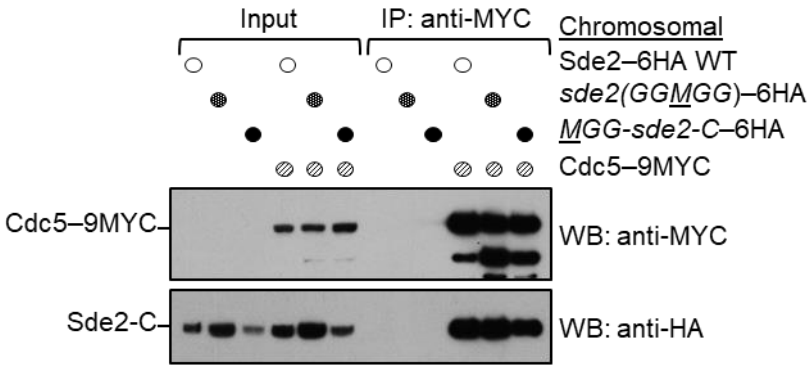


Figure 3.8.2 – CoIP assay to monitor association of ^{Met}Sde2-C with Cdc5
^{Met}Sde2-C generated from processing of *GGMGG* mutant precursor or after translation *de novo* does not show obvious defects in spliceosomal association. Sde2 mutant protein levels in both input and IP samples were variable due to difference in their stability (shown in Fig.3.8 C and Fig. 3.8.1)

3.9 Sde2-C with lysine at the N-terminus facilitates recruitment of Cay1 to the spliceosome

The protein turnover rates of lysine mutants of ^{Lys}Sde2-C did not correlate well with their ability to function like WT Sde2. Even though s*de2 GGMGG* and ^{Met}Sde2-C mutants were stable than Sde2 WT (Fig.3.8.1), they showed defective growth at 25°C and 37°C (Fig.3.8 B). However, the spliceosomal incorporation of these mutants was unaffected indicating that the lysine was not critical for the association of Sde2-C with the spliceosomes (Fig.3.8.2). So, we questioned whether lysine residue in ^{Lys}Sde2-C has any specific function in the spliceosome.

Meanwhile, studies on the effect of *sde2* deletion on spliceosome composition were under investigation, where spliceosomes from wild-type and *Δsde2* strains were purified using chromosomal Cdc5-6HA and Prp19-6HA tagged splicing factors and analyzed by mass spectrometry. We noticed that, whereas the composition of spliceosomes in the two strains was almost identical, peptides for Cactin/Cay1 were diminished in spliceosomal purifications from *Δsde2* strains. (Fig.3.9A) (performed by my colleague Kiran Kumar Kolathur).

A

No. of unique peptides in mass spectrometry					
	IP: Sde2- 6HA	IP: Cdc5-6HA		IP: Prp19-6HA	
Protein ↓	WT	WT	*Δsde2*	WT	*Δsde2*
Cdc5	61	72	74	69	69
Prp19	25	28	28	26	23
Cay1	**26**	**15**	**9**	**9**	**1**

B

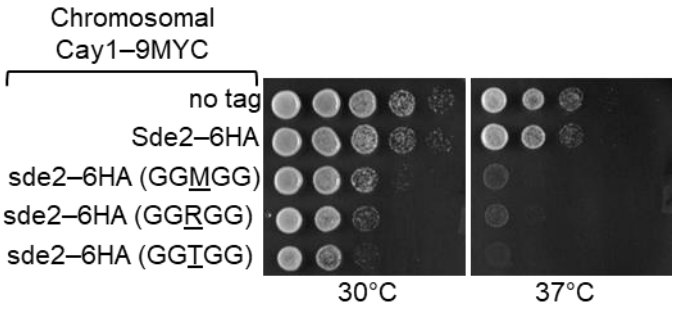


Figure 3.9 – Sde2-C facilitates association of Cay1 with spliceosomes

A. Sde2 CoIPs with Cactin in the spliceosome. Cdc5–6HA and Prp19-6HA complexes were immunoprecipitated using anti-HA antibody beads from *S. pombe* lysates wild-type *and Δsde2* strains. CoIP proteins were analyzed by mass spectrometry. The table shows number of unique peptides obtained for each protein in mass spectrometry. Reduction in number of unique peptides for Cay1 is highlighted by red coloured numbers (results from my colleague Kiran Kumar Kolathur).

B. Growth phenotypes of chromosomal Sde2 lysine mutants on rich media at 30°C and 37°C.

We tested whether or not the lysine residue in ^{Lys}Sde2-C is important for specific association with Cay1. We chromosomally mutated lysine in Sde2-6HA to methionine (^{Met}Sde2-C), arginine (^{Arg}Sde2-C) and threonine (^{Thr}Sde2-C), in chromosomal Cay1-9MYC epitope-tagged background. Growth phenotype of these strains was monitored by spot assays and incubation at 30°C and 37°C. Interestingly, all three lysine mutants were defective in growth (Fig.3.9 B), although ^{Met}Sde2-C and ^{Arg}Sde2-C showed better growth than ^{Thr}Sde2-C. Further, to monitor the association of these mutants with Cay1 in the spliceosome, Sde2 mutants were immunoprecipitated with an anti-HA antibody, and the relative association of

Cay1 was observed by anti-MYC immunoblotting.

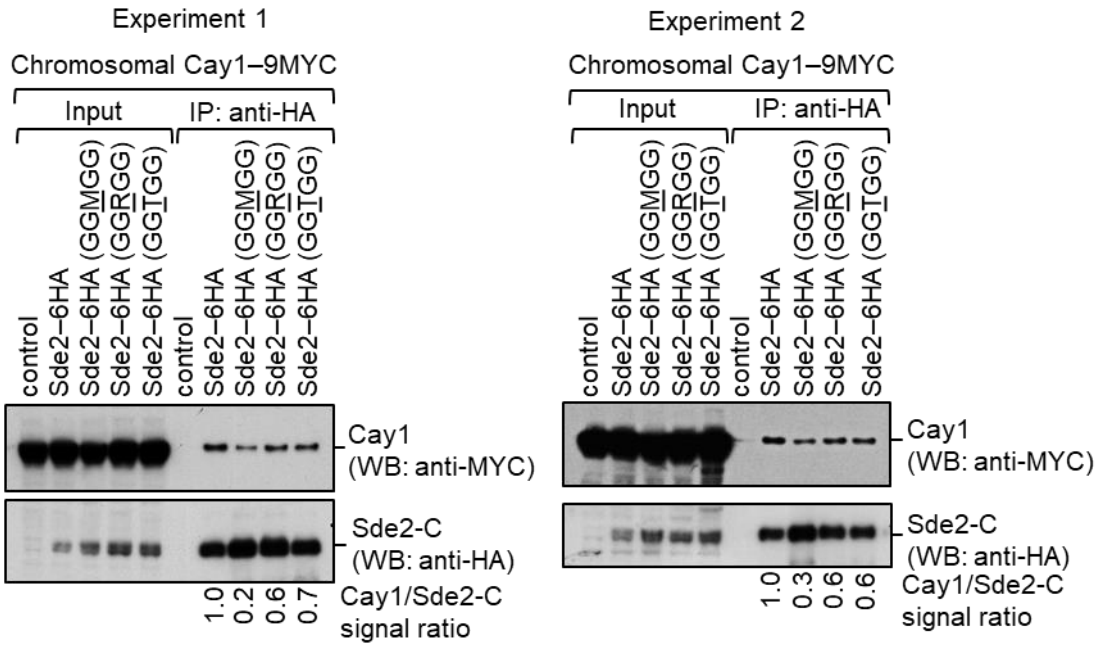


Figure 3.9.1 – N-terminal lysine of Sde2-C is crucial for interaction with Cactin
Association of Cactin/Cay1 with lysine mutants of Sde2-C. Experiments are performed with proteins expressed from chromosomally tagged strains. IP was performed using anti-HA beads followed by immunoblot assays with anti-MYC antibody to monitor Co-IP of Cay1. Anti-HA antibody was used to monitor the IP efficiency of Sde2 mutants. Numbers indicate the ratio of Cay1 to Sde2 (MYC/HA) signals obtained from ImageJ quantification of immunoblots. Results from two independent biological replicate experiments are shown.

In CoIP assay of Cay1, Sde2 mutants were immunoprecipitated with anti-HA antibody coupled beads and level of Cay1 association was monitored by anti-MYC immunoblotting. ^{Lys}Sde2 mutants were expressed at different levels because of change in their turnover rates due to their N-terminal amino acid (INPUT and IP lanes of Expt.1 and Expt.2 of Fig.3.9.1). Therefore, to compare the relative strength of Cay1 association with each Sde2 mutant, we quantified the protein bands in immunoblot using ImageJ software. The level of coIPed Cay1 with ^{Met}Sde-C was strongly diminished and with ^{Arg}Sde2-C and ^{Thr}Sde2-C was weaker than the wild type ^{Lys}Sde-C (observe lane 7,8,9,10 upper panel of IP samples in Expt.1 and Expt.2 of Fig.3.9.1). Thus, the extent of complementation of *Δsde2* by the lysine mutants appears to depend on their ability to associate with specific pre-mRNA splicing factors and turnover by the N-end rule pathway (experiments on protein turnover of ^{Lys}Sde2 mutants were done by my colleague Sumanjit Datta [200]).

3.10 ^{Lys}Sde2-C mutants show splicing defects in *ftp105*

To monitor the effect of these lysine mutants on splicing function, the total RNA was harvested after 37°C heat shock, 10 min and splicing defects in *ftp105* were checked by RT-PCR assay. Intron-3 of ftp105 was one of the top hits in pre-mRNA splicing targets of Sde2, obtained from splicing-sensitive microarray data [200]. All three lysine mutants (*GGMGG*, *GGRGG* and *GGTGG*) gave splicing defects for intron-3 of ftp105. However, the splicing defects were not identical to *Δsde2* (Fig.3.10).

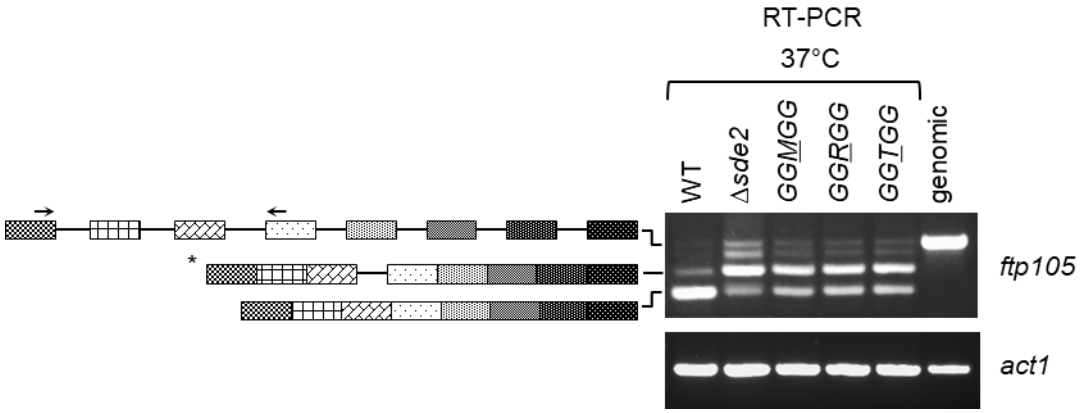


Figure 3.10 – Intron-specific splicing defects in processing defective and lysine mutants of Sde2. *S. pombe Δsde2* strain was transformed with Sde2 WT, empty vector, *sde2 GGMGG*, *sde2 GGRGG* and *sde2 GGTGG*. Black arrows marks the primers used for RT-PCR assay. Asterisk (*) marks the intron-containing transcripts accumulated in ***Δsde2, sde2 GGMGG, GGRGG and GGTGG mutant*** strains.

These results suggest that the conserved lysine may not be critical for the spliceosomal incorporation but, it is crucial for intron-specific splicing function of Sde2-C. *S. pombe* Ftp105 is a golgi localized protein implicated in vesicle-mediated transport. Ftp105 is known to control localization of Ubp5 [52] which is one of the activators of Sde2. There could be a positive feedback loop existing between Ubp5, Sde2 and Ftp105 for regulatory purpose and requires further studies.

3.11 Ubiquitin like processing of Sde2 is conserved in humans

Multiple sequence alignment of putative orthologs of Sde2 protein showed the presence of the di-glycine motif in C1orf55, the human ortholog of Sde2, hereafter referred to as *Hs*Sde2 (Fig.1.6.1). To examine evolutionary conservation of Sde2 processing, an epitope-tagged version of the *Hs*Sde2 was generated and its GGKGG motif was mutated to *AAKGG*, *GGAGG*. Similar to *S. pombe*, *Hs*Sde2 precursor was also cleaved into N-terminal ubiquitin-fold (*Hs*Sde2$_{UBL}$) and C-terminal domain (*Hs*Sde2-C) in mammalian cells (lane 1 of anti-FLAG and anti-MYC immunoblots in Fig.3.11).

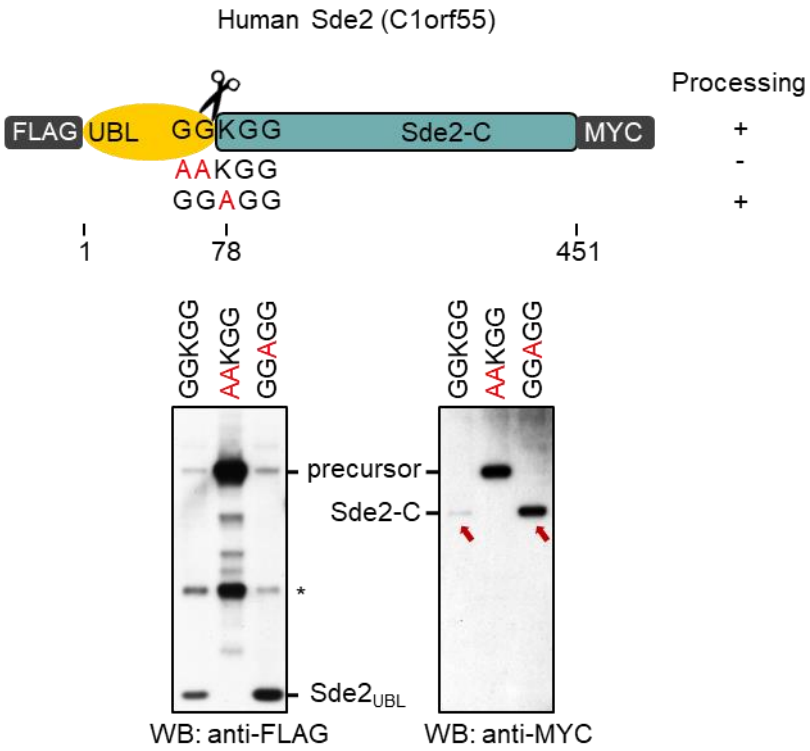


Figure 3.11 – Mammalian Sde2 is also processed after the di-glycine motif (GG~KGG) like *S. pombe* Sde2: Expression of human Sde2 (C1orf55) protein in U2OS cells. Constructs with sequences encoding 3FLAG epitope tag at the N-terminus of *HsSDE2* gene and single *MYC* epitope tag at its C-terminus under *CMV* (cytomegalovirus) promoter were used. Residues marked in red color show mutations of wild-type Sde2. Red arrows in anti-MYC blot indicate that the wild-type *Hs*Sde2-C protein having lysine at its N-terminus accumulates to a lower level as compared to *Hs*Sde2-C protein which starts with an alanine (data is from Tim Ammon and Shravan Kumar Mishra). Asterisk mark the non-specifically processed band in anti-FLAG wester blot.

The alanine mutations of the first GG residues (*AAKGG*) in *Hs*Sde2 abolished its processing (lane 2 of anti-FLAG and anti-MYC immunoblots in Fig.3.11), whereas mutation of the lysine (*GGAGG*) did not show any visible effect on processing (lane 3 of anti-FLAG and anti-MYC immunoblots in Fig.3.11). We hypothesized that *Hs*Sde2 could also be processed after the di-glycine motif like *S. pombe* Sde2 (data obtained from Tim Ammon and Shravan Kumar Mishra). Recently, the PCNA dependent processing and degradation of human Sde2 in response to replication stress was also reported [188].

The results of *Hs*Sde2 being processed similarly like *Sp*Sde2 grabbed our attention towards the conservation of its processing enzymes. The human homolog of Ubp5 and Ubp15 in humans, USP7 is known to deubiquitinate p53 and MDM2 [202]. To check whether *Hs*USP7 is capable of processing either *Hs*Sde2 or *Sp*Sde2, we co-expressed this enzyme with *Sp*Sde2, *Hs*Sde2, and Ub-Sde2-C (as a control) in *E. coli*. Strikingly, the immunoblot assays revealed that *Hs*USP7 is neither able to cleave *Sp*Sde2 nor *Hs*Sde2 but actively

processed Ub-Sde2-C *in vitro* (Fig.3.11.1 A). A parallel experiment revealed that *Sp*Ubp5 and *Sp*Ubp15 also could not process *Hs*Sde2 *in vitro* (Appendix section Fig.3.11.2).

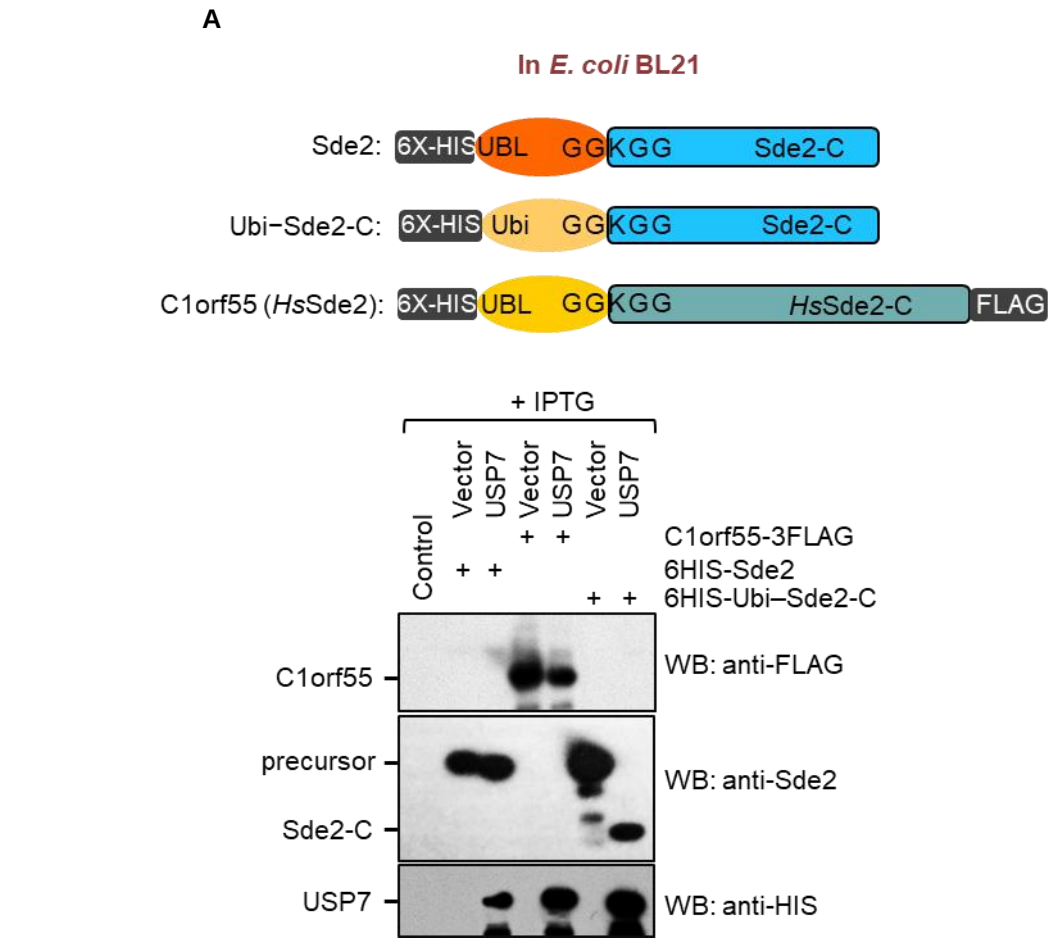


Figure 3.11.1 – Processing of *Sp*Sde2, C1orf55 (*Hs*Sde2), and Ub-Sde2-C by *Hs*USP7
A) Expression constructs harbouring indicated gene or cDNAs were co-transformed in *E. coli* BL21 (DE3) strain. Following protein expression, total cell lysates were processed by immunoblotting using anti-Sde2-C and anti-FLAG antibody. The proteases expression was checked by anti-His antibody.
B) Multiple sequence alignment (visualized using Jalview,[203]) and comparison between *S. pombe* Sde2 UBL, Ubiquitin, Ned8, and SUMO.

The amino acid sequence comparison of *S. pombe* Sde2 with ubiquitin revealed that *S. pombe* Sde2 and ubiquitin are less than 20% identical to each other (Fig.3.11.1 B) (Alignment was done using Jalview software [203] and Clustal Omega online tool). The hydrophobic patches of ubiquitin (Ile44, Ile36, and Phe4 patch) or the TEK box and Asp58 patch important for enzyme-substrate interaction were not identical to *Sp*Sde2. However, the C-terminal tail residues of ubiquitin (69-LVLRLRGG-76) and Sde2 (77-LCTRVLGG-84) appeared much similar. One of the possible mechanisms governing substrate recognition by either Ubp5 or Ubp15 could be the C-terminal tail, which in case of ubiquitin, extends in the cleft between thumb and palm sub-domains of the USP catalytic core. Importantly the ubiquitin or Sde2 binding to the finger sub-domain of USP along with their C-terminal tails might contribute to the specificity.

It is apparent from the above results that *Hs*USP7 retained its ubiquitin-specific protease activity but lost *Sp*Sde2 specific processing activity in the evolution. However, the question arises here is, what is the region of *Hs*USP7 that got evolved and led to the loss of its Sde2 processing activity but not ubiquitin processing activity.

3.12 MATH and USP domain mutations of Ubp5 and Ubp15 affect processing of *Sp*Sde2 *in vitro*

To investigate the factors responsible for dual specificity of *S. pombe* Ubp5 and Ubp15, we studied the domain architecture of these proteases. The enzymes Ubp5 (1108 aa) and Ubp15 (1129 aa) harbour MATH and USP domains followed by a C-terminal tail. The human homolog of these DUBs, USP7 is extensively studied for its domain architecture and processing activity on ubiquitin. USP7 have five UBL domains at its C-terminus in addition to the MATH and USP domains. We asked whether MATH, USP or C-terminal domains of Ubp5 and Ubp15 have any unique feature which makes them specific for Sde2. To address this question, we aligned Ubp5 and Ubp15 with USP7 to compare the amino acid sequence identity (Fig. 3.12) and also predicted the structures of Ubp5 and Ubp15 (Fig.3.12.1) using i-TASSER protein prediction program [185].

Comparison of the predicted structures of *S. pombe* Ubp5 and Ubp15 with the known structure of USP7 (Fig.3.12.1 C) [204], revealed that both *S. pombe* DUBs have similar USP and UBL domain architecture. The predicted structures showed presence of five UBL domains in Ubp15 and three UBL domains in Ubp5 at their C-termini (Fig.3.12.1 A & B). The MATH domains are not shown in the figure to avoid complexity. Therefore for comparison, USP catalytic core and UBL domains are displayed in Fig.3.12.1.

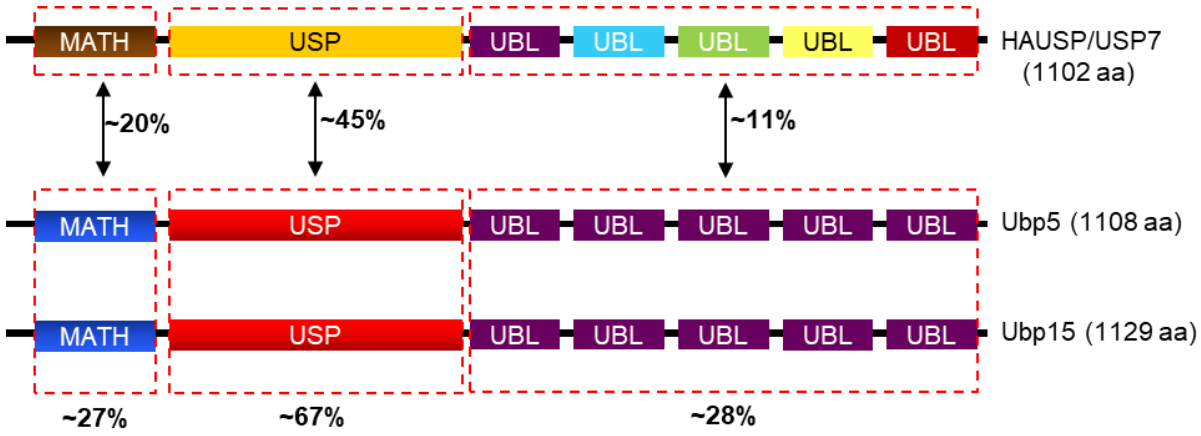
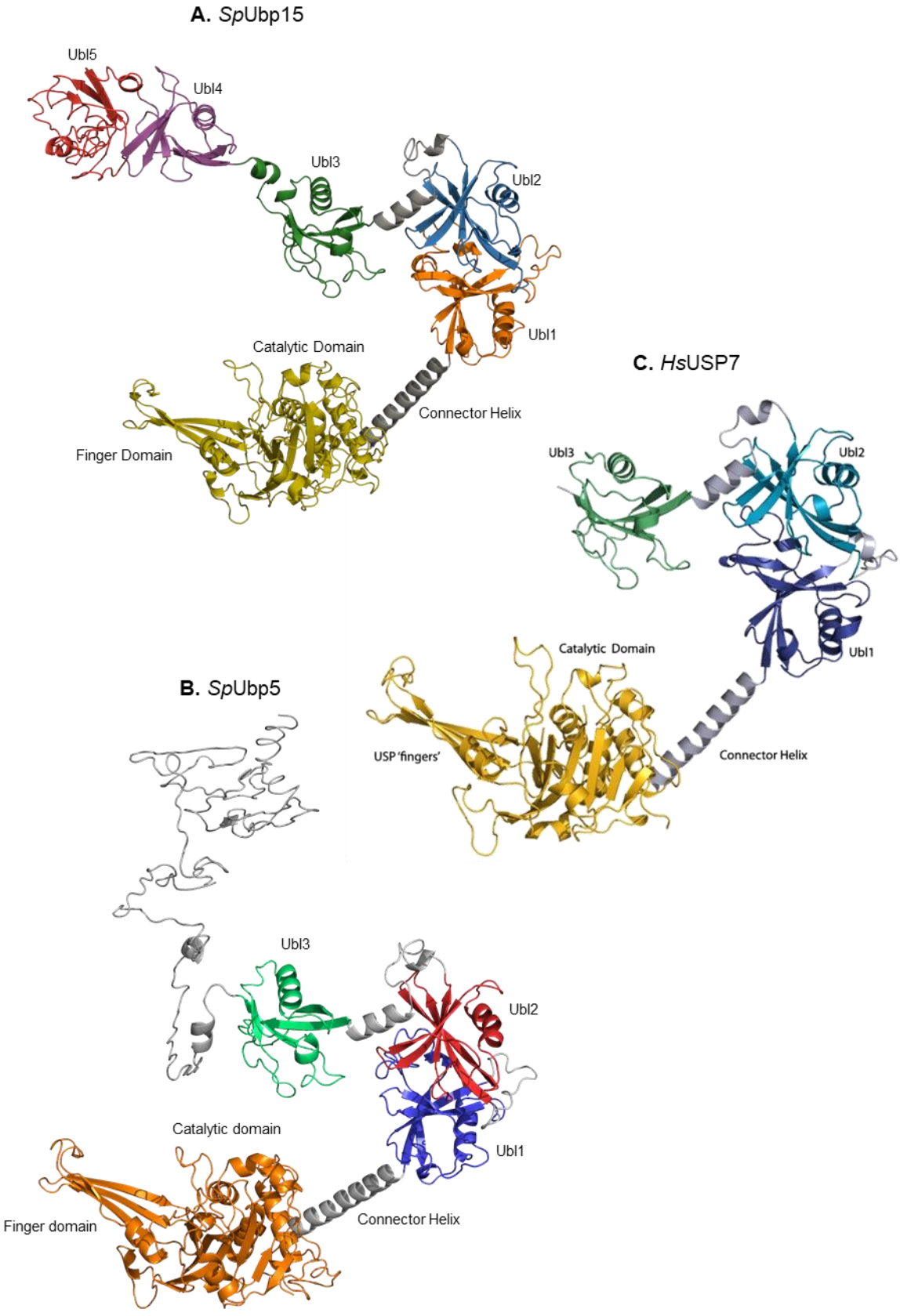


Figure 3.12 – Schematic of MATH, USP, and UBL domain sequence identity between *S. pombe* Ubp5, Ubp15 and *Hs*USP7: The percentage amino acid sequence identity between domains of *S. pombe* Ubp5 and Ubp15 is indicated below each red coloured box highlighting the domains. For *Hs*USP7 the identity with either Ubp5 or Ubp15 is shown by a two headed arrow. The length of each DUB is mentioned in the parenthesis. Sequence alignment and amino acid sequence identity of each domain was done using Clustal Omega online tool.

To study the role of sub-domains of these proteases, we made truncated versions of either *Sp*Ubp5 or *Sp*Ubp15 and checked their activity on Sde2 and ubiquitin. We also made domain swapped chimeras of *Sp*Ubp15 with *Hs*USP7 and monitored loss of protease activity on Sde2 without affecting their ubiquitin processing ability. *Sp*Ubp5, *Sp*Ubp15, and *Hs*USP7 which cleaved ubiquitin in *E. coli* were used as controls in enzyme activity assay. The list of chimeras or truncated mutants of either MATH, USP or UBL domains of these proteases are as follows:

1) First clone was full-length *Sp*Ubp15 which was active on both Sde2 and ubiquitin.
2) Second clone was full-length *Hs*USP7 which was active on ubiquitin.
3) The third chimera was made by fusion of *Hs*USP7 MATH & USP domains (1-537 aa) with *Sp*Ubp15 UBL domains (562-1129 aa). This chimera, encoded as *7USP-15UBL* (in Fig.3.12.2) was used to monitor the gain of Sde2 processing activity by *Hs*USP7.
4) In the fourth chimera, *Sp*Ubp15 MATH & USP domains (1-561 aa) was extended with 60 aa non-specific region in place of UBL domains. This chimera encoded as *ubp15ΔUBL* (in Fig.3.12.2) was used to study the role of UBL domains of *Sp*Ubp15 in Sde2 processing activity. Here, the loss of Ubp15's activity on Sde2 or ubiquitin was monitored.



5)

Figure 3.12.1 – Structural comparison between *Sp*Ubp5, *Sp*Ubp15 and *Hs*USP7: A) i-TASSER predicted structure of *S. pombe* Ubp15 showing similar domain architecture like USP7. The USP subdomains and C-terminal UBL domains are marked. **B)** i-TASSER predicted structure of *S. pombe* Ubp5, the region corresponding to UBL4 and UBL5 domain was a random coiled coli in case of Ubp5

but first three UBL domains were predicted. **C)** The crystal structure of *Hs*USP7 catalytic USP domain, connector helix and three UBL domains [204]. The structure of MATH domain is not shown here and only catalytic and UBL domains are displayed for comparison.

6) In the fifth chimera, we divided the USP catalytic domain into two halves where *Sp*Ubp15 1-379 aa region was fused to *Hs*USP7 364-1102 aa region. We wanted to study whether USP catalytic core sub-domains of either *Sp*Ubp15 or *Hs*USP7 can be swapped with each other without affecting their protease activity. The chimera was encoded as *U15U7* (in Fig.3.12.2).

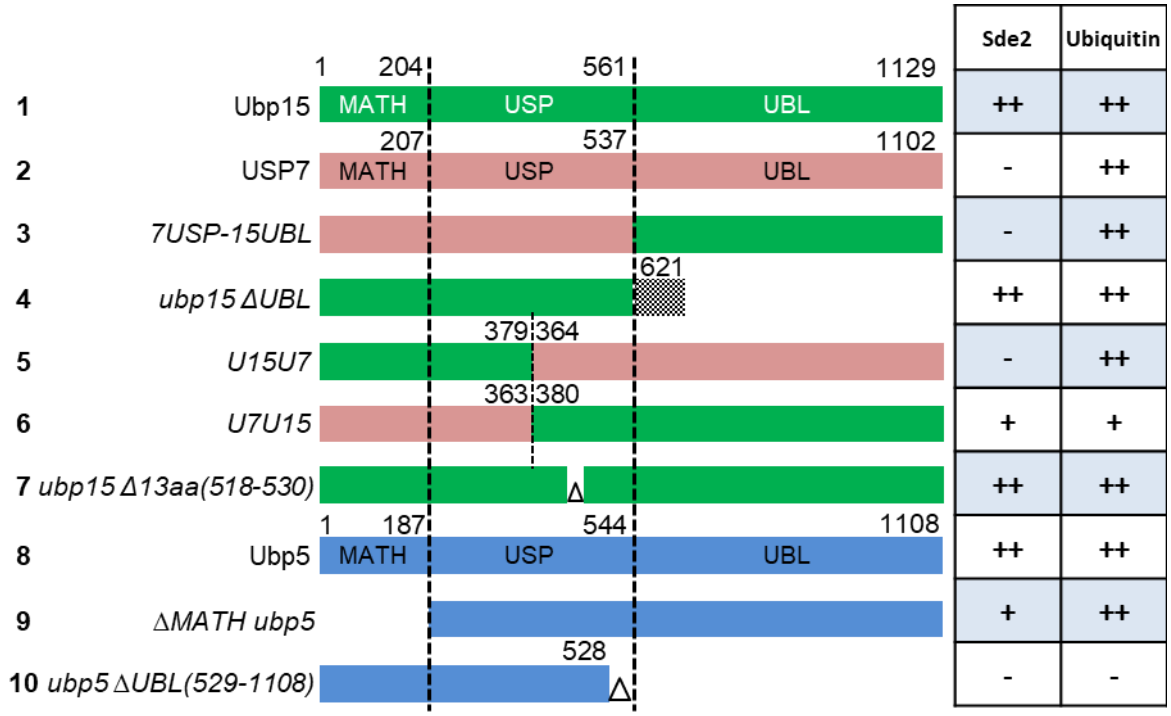


Figure 3.12.2 – Schematics of chimeras or truncated mutants of proteases used the study. The domains of Ubp5 or Ubp15 or USP7, namely MATH, USP and C-terminal UBLs is demarked by dotted vertical like. The amino acid deletion or truncation in USP domain is mentioned by Δ sign. Non-specific region extended at the C-terminus is shown with a textured grey box. Summary of processing activity of domain swapped chimeras or truncated mutants on Sde2 and ubiquitin is given in table. The activity of WT Ubp15/Ubp5 on Sde2 and Ubiquitin is shown by ++ sign. The activity of USP7 on ubiquitin is shown by ++ sign. For poor protease activity of the chimera or mutants (+) sign is used and for loss of protease activity, (–) sign is used in the table.

7) The sixth chimera contained a fusion of *Hs*USP7 1-363 aa region fused to *Sp*Ubp15 380-1129 aa region and was named as *U7U15* (in Fig.3.12.2). The approach was similar to the fifth chimera, and we wanted to narrow down the catalytic region required for Sde2 specific activity.

8) Here we made a mutant of *Sp*Ubp15 by deleting the variable region of 13 amino acids from the USP catalytic domain of *Spubp15 (Δ518-530 aa)*. The mutant was named as *ubp15 Δ13(518-530)* in Fig.3.12.2.
9) The eighth clone was full-length *Sp*Ubp5 which cleaved both Sde2 and ubiquitin.
10) In the ninth chimera, truncation mutant was made by deleting MATH domain (1-187 aa) of *Sp*Ubp5 to check the importance of the MATH domain in Sde2 specific protease activity. The mutant was named as *ΔMATH ubp5* in Fig.3.12.2.
11) In the tenth chimera, C-terminal region of USP and UBL domains (529-1108 aa) was deleted from *Sp*Ubp5. Here, in addition to UBL deletion, the USP catalytic domain was truncated. This mutant of *Sp*Ubp5 names as *ubp5 ΔUBL(529-1180)* was used to check the effect of truncation of catalytic domain on substrate processing activity (Fig.3.12.2).

The schematic of the constructed clones are shown in figure 3.12.2. The constructs were co-expressed with Sde2 and Ub-Sde2-C in *E. coli* BL21 (DE3) strain. The processing efficiency was monitored by anti-Sde2 immunoblotting and protease chimeras/mutants were detected by anti-HIS immunoblotting (Fig.3.12.3).

The deletion of UBL domains from *Sp*Ubp15 did not affect its processing activity on either substrate (chimera 4 in Fig. 3.12.2 & Fig.3.12.3), while replacement of *Hs*USP7 UBL domains with *Sp*Ubp15 UBL domains (chimera 5 or *7USP-15UBL*) could not gain the activity on Sde2, however, it efficiently processed ubiquitin. Therefore, the UBL domains of *Sp*Ubp15 are dispensable for Sde2 specific processing activity *in vitro*. The chimera 5 or *U15U7* (in Fig. 3.12.3) lost its Sde2 processing ability but it efficiently cleaved ubiquitin. Interestingly, the chimera 6 or *U7U15* was active on both Sde2 and ubiquitin, albeit to a lesser extent than the WT *Sp*Ubp15 (Fig.3.12.2 & Fig.3.12.3). The chimera 7 or *ubp15 Δ13(518-530)* where 13 aa variable region was deleted from *Sp*Ubp15 was active on both Sde2 and ubiquitin similar to its WT version (Fig.3.12.2 & Fig.3.12.3). The deletion of MATH domain from *Sp*Ubp5 also affected the processing of Sde2 however it actively processed ubiquitin (chimera 9 in Fig.3.12.2 & Fig.3.12.3). The UBL deletion construct of *Sp*Ubp5 (chimera 10 in Fig.3.12.2 & Fig.3.12.3) which also have a C-terminal truncated USP domain (deletion of 530-544aa region), lost its activity on both the substrates.

From the above results, we confirmed that the MATH domains of *Sp*Ubp5 or *Sp*Ubp15 are important for Sde2 processing activity *in vitro*. The UBL domains of *Sp*Ubp15 are dispensable for the processing of Sde2 and ubiquitin *in vitro*. The USP catalytic domains of *Sp*Ubp5 or *Sp*Ubp15 are crucial for optimum Sde2 and ubiquitin processing activity *in*

vitro. The specificity of *Sp*Ubp15 for Sde2 lies within the USP core catalytic region of 380-561 amino acids.

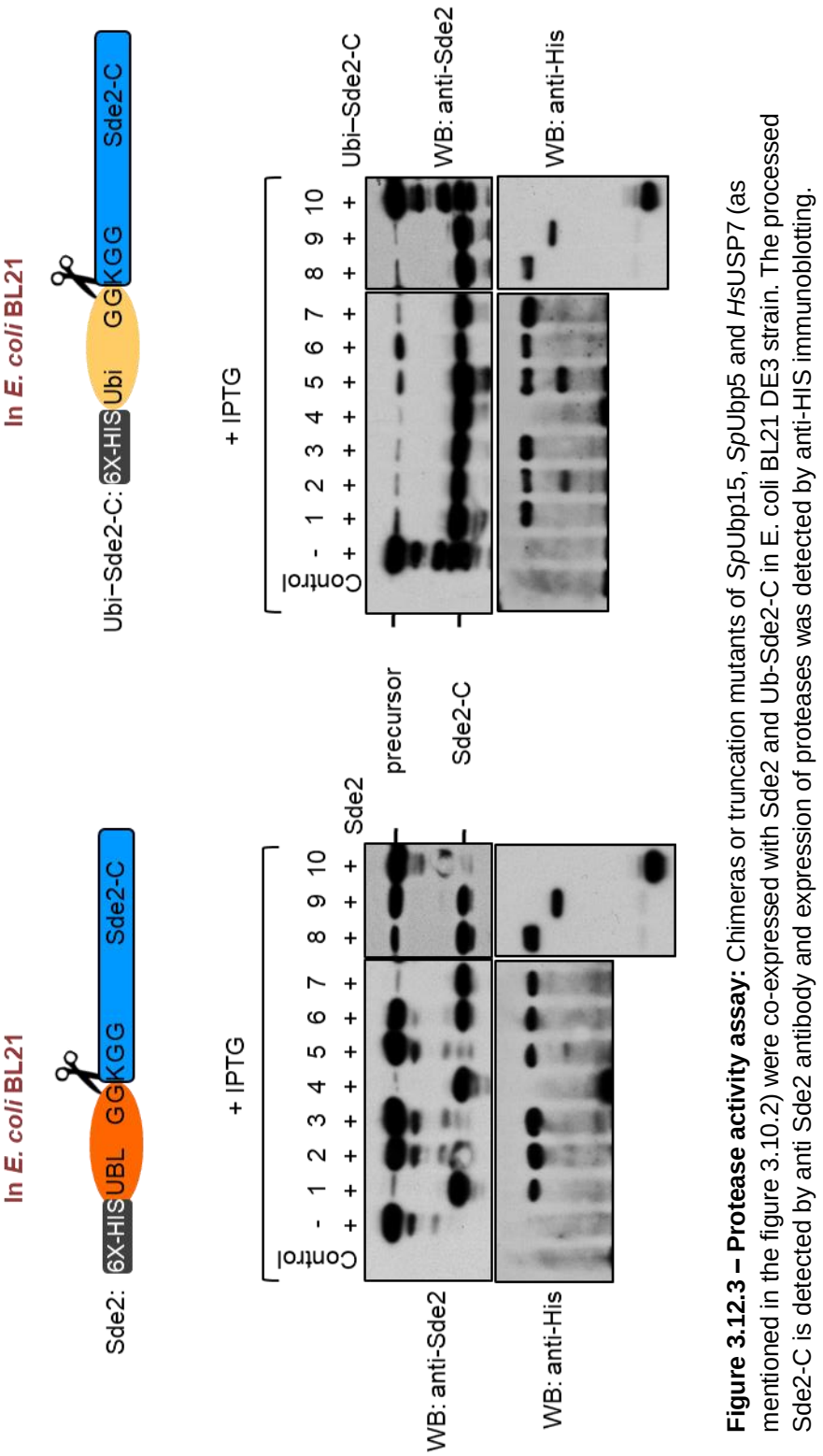


Figure 3.12.3 – Protease activity assay: Chimeras or truncation mutants of *Sp*Ubp15, *Sp*Ubp5 and *Hs*USP7 (as mentioned in the figure 3.10.2) were co-expressed with Sde2 and Ub-Sde2-C in E. coli BL21 DE3 strain. The processed Sde2-C is detected by anti Sde2 antibody and expression of proteases was detected by anti-HIS immunoblotting.

Discussion and Conclusion

4.1 Discussion

In this study, we deciphered an intriguing mechanism of activation of ubiquitin-like protein Sde2 and its role in pre-mRNA splicing. Using biochemistry, molecular biology and yeast genetics approach, we discovered that USP domain DUBs Ubp5 and Ubp15 process a ubiquitin fold containing splicing regulator Sde2 and thereby regulate intron-specific pre-mRNA splicing in *S pombe*. Sde2 is synthesized as a precursor having an N-terminal ubiquitin fold ($Sde2_{UBL}$) with a conserved di-glycine motif and a C-terminal Sde2-C domain. Sde2 is processed at the conserved GG~K site by deubiquitinating enzymes Ubp5 and Ubp15. Processing of Sde2 precursor facilitates its association with the spliceosome where $Sde2_{UBL}$ helps in the optimal expression of Sde2-C. Intriguingly, $Sde2_{UBL}$ plays an inhibitory role for incorporation Sde2-C into the spliceosome. The ^{Lys}Sde2-C also facilitates the recruitment of Cactin/Cay1 to the spliceosome and is crucial for efficient spliceosomal function. The synthetic sickness shown by *S. pombe* Δ*ubp5 ubp15* strain is a cumulative effect of impaired deubiquitinating function and intron-specific splicing defects. The phenomenon of processing of Sde2 at the GG~K site is conserved in humans. However, the human homolog of *S. pombe* Ubp5 and Ubp15, USP7 has lost the ability to process *Sp*Sde2 and *Hs*Sde2 *in vitro*. The DUBs Ubp5 and Ubp15 recognized two distinct substrates i.e., $Sde2_{UBL}$ and ubiquitin which are less than 20% identical to each other. The MATH and USP domain mutants of these DUBs affect Sde2 processing while UBL domains of Ubp15 are dispensable for processing of Sde2 and ubiquitin.

4.1.1 Sde2 is a new member of the UBL family

Sde2 protein is conserved from fission yeast to humans. Multiple lines of evidence confirmed that Sde2 is a *bona fide* UBL. First, $Sde2_{UBL}$ harbour a predicted ubiquitin fold or the β-grasp fold found in UBL family of proteins. Second, the $Sde2_{UBLGG}$–Sde2-C precursor is cleaved at the di-glycine motif by ubiquitin-specific proteases Ubp5 and Ubp15. Third, $Sde2_{UBLGG}$–GFP fusions were processed like Sde2 precursor. Intriguingly the amino acid sequence homology between $Sde2_{UBL}$ and ubiquitin is ~12%, which makes them distinct substrates and yet they are recognized by the same enzymes.

UBLs are often synthesized as precursors which get processed after the di-glycine motif by the DUBs or UBL-specific proteases. In case of Sde2, the processing resulted in the

generation of two domains, the N-terminal $Sde2_{UBL}$ and a C-terminal Sde2-C. Complementation assays with processing defective and *de novo* expressed versions of Sde2 demonstrated that expression of $Sde2_{UBL}$ could not rescue the growth defects of *Δsde2* strain. However, the expression of Sde2-C partially complemented the growth defects indicating that Sde2-C is the functional domain. Deletion of $Sde2_{UBL}$ coding region from chromosomal locus affected the protein level of Sde2-C and showed severe growth defects. We also found that chromosomal replacement of $Sde2_{UBL}$ with ubiquitin could complement Sde2's function (data not shown). Presumably the function of $Sde2_{UBL}$ is to regulate the optimal expression and folding of Sde2-C similar to ubiquitin's chaperone-like activity reported for the proper expression of ribosomal subunits from the ubiquitin ribosomal precursors [3,6] One of the main functions of $Sde2_{UBL}$ appears to generate the functional Sde2-C starting with a lysine, as the mutants with N-terminal methionine, arginine or threonine showed growth defects and *Δsde2*-like intron-specific splicing defects. Thus, altogether, the $Sde2_{UBL}$ and its processing are absolutely critical to facilitate the generation of spliceosomal Sde2-C with lysine at its amino-terminus which is crucial for intron-specific splicing activity.

A distinguishing feature of Sde2 homologs is the conserved GGKGG motif which is critical for the processing and function. The human ortholog of Sde2, C1orf55 (herein referred to as *Hs*Sde2) is also synthesized as a ubiquitin fold containing precursor harboring the conserved diglycine motif. Similar to *Sp*Sde2, the *Hs*Sde2 precursor is processed after the first di-glycine motif to generate *Hs*Sde2-C starting with a lysine. Additionally, both *Sp*Sde2 and *Hs*Sde2 share moderate sequence similarity. Recently, PCNA-dependent cleavage and degradation of *Hs*Sde2 is shown to regulate replication stress [188]. Thus, both *Sp*Sde2 and *Hs*Sde2 are processed like ubiquitin-precursors but still seem to retain organism-specific features.

4.1.2 ^{Lys}Sde2-C is a unique pre-mRNA splicing regulator

^{Lys}Sde2-C appears to function differently from other pre-mRNA splicing regulators of the spliceosome. ^{Lys}Sde2-C promotes efficient excision of selected introns from selected transcripts in *S. pombe*, but it is not required for general pre-mRNA splicing. ^{Lys}Sde2-C thereby becomes a critical control factor for the expression of selected proteins, a majority of which function at the chromatin. Therefore, growth defects or drug sensitivities of Δ*sde2* strain could not be attributed to splicing defects of individual genes. The lack of any obvious common feature in Sde2 target pre-mRNAs leads us to postulate that some RNA secondary structures or length of introns or distance between the BS and 3' ss could make

their splicing Sde2 dependent [200]. Importantly, Sde2-specific proteases play a critical regulatory role in pre-mRNA splicing by processing the inactive Sde2 precursor to generate the active spliceosomal ^{Lys}Sde2-C. The processing-deficient and lysine mutants of Sde2 display Δ*sde2*-like pre-mRNA splicing defect.

The splicing factor Cdc5 was one of the top hits in Sde2 co-purified spliceosomal proteins (Appendix Table 5.2). *S. pombe cdc5* was first identified in a screen of fission yeast mutants defective for cell cycle progression [205]. Subsequent studies have shown that *S. pombe* Cdc5 is an essential member of the spliceosome and functions in pre-mRNA splicing [206]. Cdc5 has N-terminal nucleic acid binding domain or Myb (Myeloblastosis) repeats and shows direct interaction with other core NTC/Prp19 complex members. Therefore, it is predicted that Cdc5 may facilitate NTC-mediated RNA− RNA and/or RNA−protein transitions by acting as a scaffold linking NTC components and RNAs [207]. NTC joins the spliceosome before, or during unwinding of U4 from U6 and stay associated with the spliceosome during the two steps of splicing and plays an important role in regulating spliceosome conformations and fidelity [208].

Activated ^{Lys}Sde2-C facilitates association of spliceosomes with Cactin/Cay1, a spliceosomal factor with Sde2-like function in intron-specific pre-mRNA splicing [200,209]. While the lysine does not appear crucial for the interaction of ^{Lys}Sde2-C with spliceosomes, the residue is important for interactions with Cactin. The free lysine in ^{Lys}Sde2-C also plays a regulatory role, as it makes the protein a natural substrate of the N-end rule pathway of proteasomal degradation [200,201]. Role of *S. pombe* Cactin for splicing of *rap1* pre-mRNA was reported previously [209]. Human Cactin functions in pre-mRNA splicing through interactions with splicing factors DHX8 and SRRM2 [210]. Cactin orthologs were detected in spliceosomal purifications [186,211], and it also appears to have mRNAs binding ability [212]. Cactin possibly recognizes selected pre-mRNAs and facilitates their splicing with Sde2-C containing spliceosomes. The lysine has been selected for optimum function and interaction of Sde2-C in the spliceosome, as proteins neither with stabilizing residues nor with more destabilizing residues performed better than the wild type [200]. To fine balance ^{Lys}Sde2-C level, the proteasome likely clears off non-spliceosomal pool of the protein by N-end rule pathway.

4.1.3 The feedback loop between ubiquitin-specific proteases Ubp5/Ubp15 and Sde2

Ubp5 and Ubp15 shows a dynamic localization patterns and found at the Golgi, cytosol and nucleus [52]. Ubp5 localization to the Golgi apparatus depends on Ftp105, a

Golgi localized protein implicated in vesicle-mediated transport. The deletion of *ftp105* gene disrupts Ubp5 localization from golgi to nucleus and it is speculated that excess levels of Ubp5 are targeted to the nucleus by an unknown mechanism [52]. Intriguingly *ftp105* is one of the splicing targets of Sde2. There may be a feedback loop existing between Ubp5, Sde2, and Ftp105 since Ubp5 activates Sde2, its activation helps in splicing and expression of Ftp105 which in turn controls Ubp5's localization patterns. Ubp5 and Ubp15 are shown to be involved in membrane trafficking where deletion of all five DUBs, i.e., *ubp4, ubp5, ubp9, ubp15, and sst2* but not loss of any combination of four, shows defects in abscission, septation, polarized growth, and it also impacted cell viability in the fission yeast [213]. These DUBs potentially have both unique and overlapping substrates which could be a key reason for their degeneracy [213].

Importantly in our studies, we found that *S. pombe* Δ*ubp5* Δ*ubp15* strain shows growth defects and is sensitive to high or low temperatures. The double deletion led to abolished processing of Sde2 and showed splicing defects in specific introns of selective genes. Since Sde2 is processed in the nucleus, the nuclear pool of these proteases would be more relevant for the activation of Sde2. The negative genetic interaction between Ubp5 and Ubp15 [198] can be explained with two reasons. First, they regulate the function of two distinct substrates, i.e., Sde2 which functions in pre-mRNA splicing and ubiquitin which is part of the ubiquitin system. The second reason is that the DUBs Ubp5 and Ubp15 have both unique as well as overlapping ubiquitinated targets reported in *S. pombe* [213]. Obviously the deletion of both the proteases will elevate the accumulation of their target substrates. Therefore the combined effect of impairment of splicing function and affected protein turnover in *S. pombe* Δ*ubp5* Δ*ubp15* resulted in severe growth defects.

4.1.4 MATH and UBL domains of Ubp5 and Ubp15 may regulate the enzymes activity for Sde2

The paralogs Ubp5 and Ubp15 are the two MATH/TRAF domain-containing ubiquitin specific proteases in *S. pombe*, which were reported to process ubiquitin. We predicted the presence of UBL domains at the C-terminus of Ubp5 and Ubp15. The human ortholog of these DUBs, USP7/HAUSP, was reported to deubiquitinate and stabilize the tumour suppressor p53 [214] and its E3 ubiquitin ligase Mdm2 [202,215,216]. Indeed, USP7 was originally identified by its interaction with ICP10 protein from herpes simplex virus [217]. It also contains a MATH domain at N-terminus and five UBL domains at C-terminus.

The MATH domain is a conserved 180 aa region shared by the functionally unrelated extracellular meprins (metalloproteases) and the intracellular TNF receptor associated factor TRAF proteins [218]. The MATH domain is a fold of seven anti-parallel β-helices which function in oligomerisation or protein-protein interactions crucial to form singling complexes with TNF receptor-1. Interestingly, USP7's MATH domain accommodate the MATH domain binding-motifs of p53, Mdm2 [219] and Epstein-Barr virus nuclear antigen 1 (EBNA1) in a shallow surface groove in the middle of the β-sandwich [220]. Since p53, Mdm2 and EBNA1 are found only in vertebrates; it seems likely that they have evolved to compete with USP7 binding. Considering the high conservation of MATH domain in evolution [218], it should be expected that primordial interactors of this domain might be well conserved during evolution. It is noteworthy that MATH domains of *Sp*Ubp5 or *Sp*Ubp15 show 20% identity with that of USP7's MATH domain. Deletion of MATH domain from Ubp5 affected its Sde2 processing activity but not for ubiquitin. It has been demonstrated that, *S. cerevisiae* Ubp15's enzymatic activity on ubiquitin is regulated by N- and C-terminal domains [221] which correspond to MATH and UBL domains of S. pombe Ubp5/Ubp15 or human USP7. Removal of the N-terminal MATH/TRAF-like domain (amino acids 1-195) from S. cerevisiae Ubp15 resulted in a 2-fold decrease in catalytic efficiency [221]. Yet more studies of MATH domain interactome in *S. pombe* and mechanism of enzyme activation will uncover its regulatory role for function of Ubp5 or Ubp15.

Several ubiquitin-like (UBL) domain containing proteins are encoded in humans and they are involved in a variety of cellular functions including ubiquitin system in eukaryotic cells. The yeast proteins Rad23 and Dsk2 are the best characterized members of this group. Rad23 and its human homologue (HHR23) interacts directly with the 26S proteasome *in vitro* [222]. These proteins are proposed to function as substrate shuttles, transporting ubiquitylated proteins in the cytosol from E3 enzymes to the 26S proteasome [223]. They are redundant in nature and target multiple substrates, although in some cases show relative substrate specificity [224]. The DUBs Ubp5 and Ubp15 also show predicted UBL domains. Considering their dynamic localization and presence of multiple targets in *S. pombe*, the UBL domains might be crucial for DUBs localization to cell compartments or to the proteasome. They may also help the proteases to interact with specific substrates or help in attaining catalytically active conformation. The nuclear processing of Sde2 by these DUBs must be a regulated process i.e., how these proteases localize near Sde2 for processing and whether directed protein assembly via MATH or UBL domains is necessary for this function requires further investigation.

An interesting interplay between DUBs and E3 ligases is reported where the interaction takes place via UBL domain-containing region. A RING finger-type ubiquitin E3 ligase, BRAP (BRCA1-associated protein) undergoes auto-ubiquitylation as a result of binding to activated RAS GTPase, involved in MAPK signaling [225]. The DUBs, USP4, and USP15 are sole interactors of BRAP and have a structure featuring an N-terminal DUSP (domain in USPs) and UBL domains. USP15 and USP4 mediate the reversal of the autoubiquitylation of BRAP, whereas BRAP promotes the ubiquitylation of USP15. USP15 depletion destabilizes BRAP by promoting its proteasomal degradation. The DUBs-BRAP interaction is mediated through N-terminal DUSP-UBL domains of USP15/USP4 with the C-terminal coiled-coil region of BRAP [226]. Secondly, the binding of DUSP–Ubl domain of USP4 promotes a change of a switching loop near the active site leading to enhanced ubiquitin dissociation and efficient turnover [227]. The RING E3 ubiquitin ligase UHRF1 controls DNA methylation through ubiquitylation of histone H3 thereby recruiting DNA methyltransferase DNMT1 to newly replicated chromatin. Here the UBL domain of UHRF1 participates in structural rearrangements of UHRF1 upon binding to chromatin. Thus UBL plays an essential role in the inheritance of DNA methylation [228]. Since reversal of ubiquitin conjugation of protein substrates is one of the functions of DUBs, the domains in Ubp5 or Ubp15 may help in interaction with specific E3 ligases thereby regulating ubiquitin conjugation.

USP7'S UBL-45 region, more specifically the C-terminal 19 amino acids (residues 1084–1102), binds to a 'switching' loop present in its catalytic domain and activates USP7 to promote ubiquitin binding. This results in increased activity of USP7 by 100-fold [229]. The activation can be allosterically enhanced by the metabolic enzyme GMPS that binds to the first three UBL domains UBL-123. The binding hyper-activates USP7 by stabilization of the HUBL-45-dependent active state [230]. Thus, both the substrate recognition and deubiquitinase activity is highly regulated by catalytic and non-catalytic domains of USP7. The roles of non-catalytic domains in regulation and localization of *S. pombe* Ubp5 or Ubp15 could not be ruled out and requires further studies.

4.1.5 The USP catalytic domain of Ubp15 governs Sde2 specific processing activity

Our enzyme activity assays revealed that Ubp15 enzyme cleaved Sde2 and Ubiquitin irrespective of their UBLs or MATH domains. Although the deletion of MATH from Ubp5 and chimera 6 of USP7(1-363 aa)-Ubp15(380-1129 aa) led to diminished activity on Sde2 or both the substrates respectively. But these protease mutants/chimeras were still active. Also

the truncation mutation (chimera 10) in catalytic domain of Ubp5 completely abolished its enzyme activity. Interestingly, the swapping of 380-561 aa region from the catalytic core of Ubp15 (chimera 5) resulted in loss of only Sde2 processing activity. These results point towards the core USP catalytic domain (380-561 aa) of Ubp15 which is governing Sde2 specific processing activity.

The catalytic domain of USP7 exists in an auto-inhibited state in absence of substrate. The binding of ubiquitin initiates a series of conformational changes in the catalytic domain and also the C-terminus of the UBL domains. These conformational changes expose the activation cleft, which allows binding of C-terminal peptide of USP7, stabilizing the active conformation and the catalytic triad [229]. In case of Ubp5 and Ubp15, the mechanism by which their catalytic domains accommodate two distinct substrates i.e. Sde2 and ubiquitin needs to be further investigated.

Multiple sequence alignment of all human USPs indicates that only USP47 and USP40 contain switching loop consensus motifs that are similar to USP7 [229]. Interestingly, both the proteases harbour UBL domains similar to Ubp5/Ubp15 and their amino acid sequence identity with Ubp15 is ~21% for USP40 and ~34% for USP47. Thus we suspect that one of the two proteases could be the processing enzyme of *HsSde2* but requires *in vitro* validation.

The USP domain protease CYLD functions in selective disassembly of Lys63-linked di-ubiquitin molecules. Indeed the specificity determinants of CYLD for the isopeptide linkage are conferred by structural elements of the USP domain in the vicinity of the catalytic site. One of the ubiquitin binding loop called BL1 in CYLD USP domain (harbouring Phe759 equivalent to the ubiquitin-interacting residue Phe409 of HAUSP), undergoes a conformational change to engage ubiquitin [231].

Literature survey revealed that there are proteases categorized in deubiquitinating enzyme family but show affinity toward UBLs. Several USP DUBs can form covalent adducts with ISG15-vinyl sulfone probes, for example; USP18, the mammalian ISG15 specific isopeptidase [232]. The hydrophobic patches identified in USP18 interact with a unique hydrophobic region of ISG15. Among two ubiquitin-like domains of ISG15, the C-terminal domain is recognized and essential for USP18 activity [233]. Another enzyme USP21 showed 80-fold higher activity on ubiquitin as compared to ISG15. For USP21 activity, the interaction between Arg-72 of ubiquitin (or Arg-153 of ISG15) and GLU-304 of USP21 (residue is conserved among all active USP domains) is essential for processing both ubiquitin and ISG15 modifications [234]. However, the physiological relevance of

deISGylation activity by DUBs is poorly understood. Few more examples of cross-specificity are JAMM domain containing COP9 signalosome-CSN5 subunit. This enzyme cleave the isopeptide linkage between NEDD8 and cullin-RING E3 ligase (CRL) scaffolds [38]. The protease USPL-1 was revealed to be an efficient SUMO isopeptidase and possess SUMO C-terminal hydrolase activity, but it neither binds nor cleaves ubiquitin bound substrates [235]. Some pathogenic viruses and bacteria encode enzymes which represent cross-specific DUBs that help impair immune surveillance mechanisms that rely on ubiquitin and ubiquitin-like modifications [43,236].

4.1.6 Splicing regulation by DUBs Ubp5 and Ubp15

DUBs are functinally linked to diverse cellular processes, including protein turnover, transcription, RNA processing, DNA damage, etc. Due to their dynamic localization, structural diversity and overlapping targets it becomes challenging to assign DUBs with a specific pathway. Since these DUBs interact transiently with the substrates, scientists have used suicide substrates like ubiquitin-vinyl methyl sulfone [237]. In our studies, Ubp5 and Ubp15 could not be detected in spliceosomal purification.

Our findings have revealed that DUBs from ubiquitin system can be used to control splicing through activation of a ubiquitin fold containing splicing regulator Sde2. The double deletion of Ubp5 and Ubp15 resulted in the complete accumulation of Sde2 precursor and showed *Δsde2* like splicing defects. These two DUBs are assigned with the specific function to process/activate Sde2, an intron-specific splicing factor.

There are some questions to be addressed in future like, what is the biological relevance of Ubp5 and Ubp15's localization at the Golgi, cytosol and nucleus? How are they regulated by Ftp105 and what are the sites of interaction between these proteases and Ftp105? From our splicing assays, we know that Sde2 is required for splicing of the intron-3 of *ftp105*. It will be interesting to know that how Sde2 carries out intron specific splicing of intron-3 of *ftp105* and what other splicing factors are involved in this mechanism. We have narrowed down the region in Ubp15 governing the specificity for Sde2, but further work is needed at residue level that would show the specificity of USP domain. The structural studies of either Ubp5 or Ubp15 in complex with Sde2 or ubiquitin will be required to know the structural plasticity of these proteases.

^{Lys}Sde2-C is a substrate of the N-end rule pathway of degradation [200], and therefore it is a short-lived protein. The higher molecular weight adducts of Sde2 were observed and were confirmed to be ubiquitin conjugates of Sde2 by pull-down assays of the conjugates

under denaturing conditions [200]. It is plausible that when Ubp5 and Ubp15 process Sde2 precursor, they might also protect Sde2-C from degradation by removal of ubiquitin and allowing it to associate with the spliceosome stably.

4.2 Conclusion

In conclusion, the ubiquitin-like processing of $Sde2_{UBL}$ generates the activated ^{Lys}Sde2-C protein, which becomes functional in the spliceosome and plays a critical and selective role in pre-mRNA splicing. The processing is carried out by two USP family deubiquitinating enzymes Ubp5 and Ubp15 in *S. pombe*. $Sde2_{UBL}$ is inhibitory for incorporation of Sde2-C into the spliceosome where the UBL fold at the N-terminus helps in the optimal expression of Sde2-C. After processing Sde2-C interacts with core spliceosomal subunits like Cdc5 and Cwf21. Importantly, Sde2 functions as a unique regulator of spliceosomes where it promotes efficient excision of only selected introns from a subset of genes. The splicing targets of Sde2 mainly work in transcription, replication, telomeric silencing, and chromosome stability. Sde2-dependent intron-specific splicing appears to be a crucial factor for proper expression of Ftp105, which regulates the localization of Ubp5 to golgi complex. ^{Lys}Sde2-C generated by processing after $Sde2_{UBLGG}$ is important for recruitment of Cactin/Cay1 to the spliceosome and intron-specific splicing. The schematic of the mechanism of activation of Sde2 is shown in Fig.4.2.

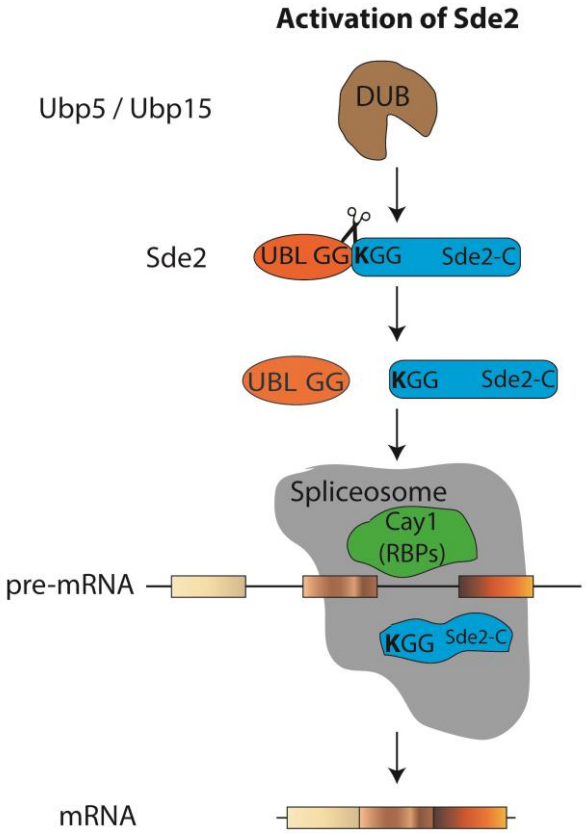


Figure 4.2 - Schematics of Sde2 activation by Ubp5 and Ubp15
The DUBs Ubp5 and Ubp15 cleave Sde2 precursor with a ubiquitin-fold (UBL) and Sde2-C domain at the GG~KGG motif. After processing the activated ^{Lys}Sde2-C becomes part of the spliceosome and facilitates recruitment of Cactin/Cay1 to the spliceosome (findings made together with my colleagues: Kiran Kumar Kolathur, Poonam Thakran and Sumanjit Datta).

We have revealed one of the intriguing properties of USP domain proteases Ubp5 and Ubp15 which are dual specific and able to process both $Sde2_{UBL}$ and ubiquitin, the substrates with less than 20% identity to each other. The human homolog of Ubp5 and Ubp15, USP7 has lost the ability to process *Sp*Sde2 and *Hs*Sde2 *in vitro*. The Sde2 specificity of Ubp15 lies within 380-561 amino acids region of USP catalytic domain.

The occurrence of precursors with ubiquitin-folds appears to be a conserved principle for both the ribosome and the spliceosome – the two major RNP complexes known to have certain similarities in their mechanisms of action [238,239]. Extending our current findings with Sde2, it is tempting to postulate that ubiquitin processing activates the ribosomal proteins for translation of selective mRNAs. Thus, we have demonstrated that two DUB paralogs Ubp5 and Ubp15 activate distinct UBLs with roles in diverse processes related to the ubiquitin system and pre-mRNA splicing.

Appendix

Table 5.1: Top pre-mRNA splicing targets of Sde2 showing affected introns (Log2 values of intron-retention ratio of *Δsde2*/WT is averaged for 4 samples: Samples were harvested at 30°C and 37°C for splicing sensitive microarrays; ns, no signal). The experiment was performed by Shravan Kumar Mishra in collaboration with Jeffrey A. Pleiss at Cornell University.

Gene Name (Systematic name)	Biological process (from PomBase)	Intron number	Log_2 intron retention ratio of *Δsde2*/WT
SPAC212.06C	DNA helicase in rearranged telomeric region	1	2.476
paa1 (SPAP8A3.09C)	Protein phosphatase regulatory subumit, cytokinesis, establishment or maintenance of cytoskeleton polarity	1	1.527
		2	0.422
		3	0.061
		4	-0.026
		5	ns
		6	-0.156
ptl2 (SPAC31G5.20C)	Triacylglycerol lipase, triglyceride mobilization	1	1.120
		2	0.476
		3	-0.085
SPBC660.16	Phosphogluconate dehydrogenase, D-gluconate metabolism, pentose-phosphate shunt	1	ns
		2	ns
		3	0.695
		4	1.479
SPAC56E4.08C	Transcription regulation	1	1.435
rap1 (SPBC1778.02)	Telomere binding protein, transcription, chromosome segregation, DNA replication	1	ns
		2	1.388
SPBC354.07C	Ergosterol biosynthesis	1	0.640
		2	1.374
		3	0.381

		4	Ns
psf3 (SPAC227.16C)	DNA replication, GINS complex subunit	1	0.148
		2	0.266
		3	ns
		4	1.344
gga22 (SPBC25H2.16C)	Golgi localized Arf binding protein, golgi to vacuole transport	1	-0.075
		2	1.300
		3	0.650
		4	0.061
		5	0.016
SPBC16E9.15	Heat shock factor binding protein	1	0.136
		2	0.072
		3	1.264
vma5 (SPAPB2B4.05)	V-type ATPase V1 subunit, proton transport	1	ns
		2	0.447
		3	1.201
		4	ns
		5	0.016
		6	ns
mcs2 (SPBP16F5.02)	TFIIH complex cyclin, transcription, cell cycle	1	-0.245
		2	1.196
apc10 (SPBC1E8.06)	Anaphase promoting complex subunit, chromosome segregation	1	1.188
		2	1.009
pdx1 (SPCC1259.09C)	Pyruvate dehydrogenase component, acetyl-CoA biosynthetic from pyruvate	1	1.176
		2	0.132
		3	0.476

		4	ns
		5	-0.060
ftp105 (SPAC17A5.16)	Ubp5 interacting protein	1	ns
		2	-0.014
		3	1.142
		4	ns
		5	ns
		6	ns
		7	-0.050
hif2 (SPCC1235.09)	Set3 complex subunit, transcription and heterochromatin assembly	1	1.104
		2	ns
		3	ns
SPAC16A10.03C	E3 ubiquitin ligase, protein transport, vesicle docking	1	ns
		2	1.090
		3	ns
rxt2 (SPBC428.06C)	Histone deacetylase complex, cell cycle, chromatin organization, transcription	1	-0.010
		2	ns
		3	0.180
		4	1.070
pre10 (SPCC1795.04C)	20 S proteasome complex subunit alpha 7, chromosome segregation	1	ns
		2	0.315
		3	1.068
ypt5 (SPAC6F6.15)	GTPase, endocytosis	1	0.188
		2	-0.050
		3	1.062
		4	ns
		5	ns

		6	-0.003
		7	ns
naa20 (SPCC16C4.12)	NatB N-acetyltransferase complex catalytic subunit, cytokinesis	1	1.054
tbp1 (SPAC29E6.08)	TATA binding protein, transcription	1	0.771
		2	1.047
		3	0.365
plp2 (SPBC2A9.09)	Phosducin family protein, transcription, protein folding	1	ns
		2	1.035
SPBC21C3.03	ABC1 kinase family protein, mitochondrial membrane kinase	1	1.020
kap114 (SPAC22H10.03C)	GTP binding karyopherin, nucleocytoplasmic transport	1	1.027
		2	-0.020
		3	0.180
		4	0.157
		5	0.120
		6	0.213
eaf1 (SPCC1223.10C)	RNA polymerase II transcription elongation factor	1	1.011
		2	ns
		3	-1.122
ubc11 (SPCC1259.15C)	E2 ubiquitin conjugating enzyme, chromosome segregation	1	-0.250
		2	1.009
		3	0.070
		4	0.068
dcd1 (SPBC2G2.13C)	Deoxycytidine deaminase, cell cycle	1	1.006
		2	-0.159
		3	-0.040

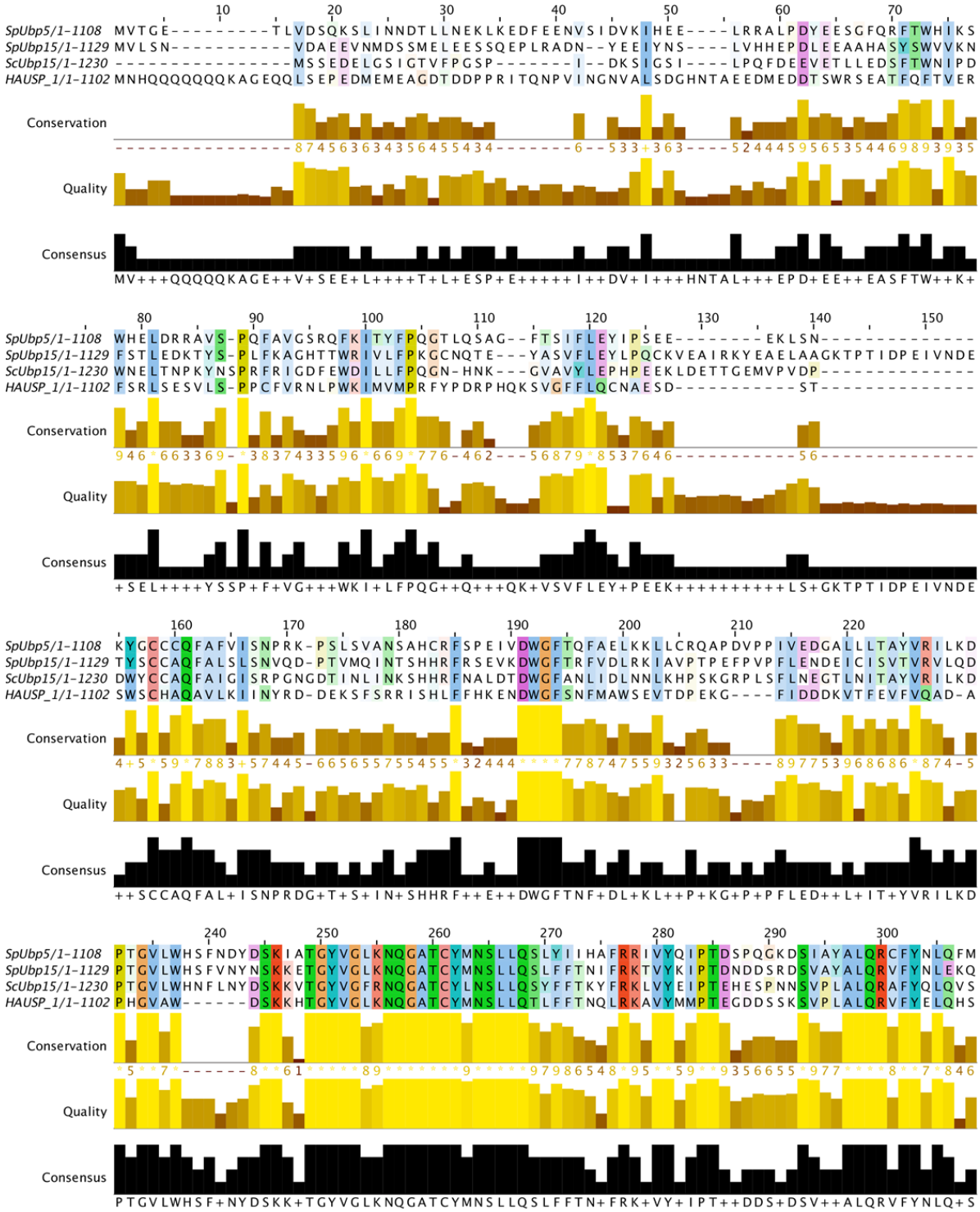
Table 5.2 (Related to figure 3.1) Sde2 co-purifies spliceosomal proteins

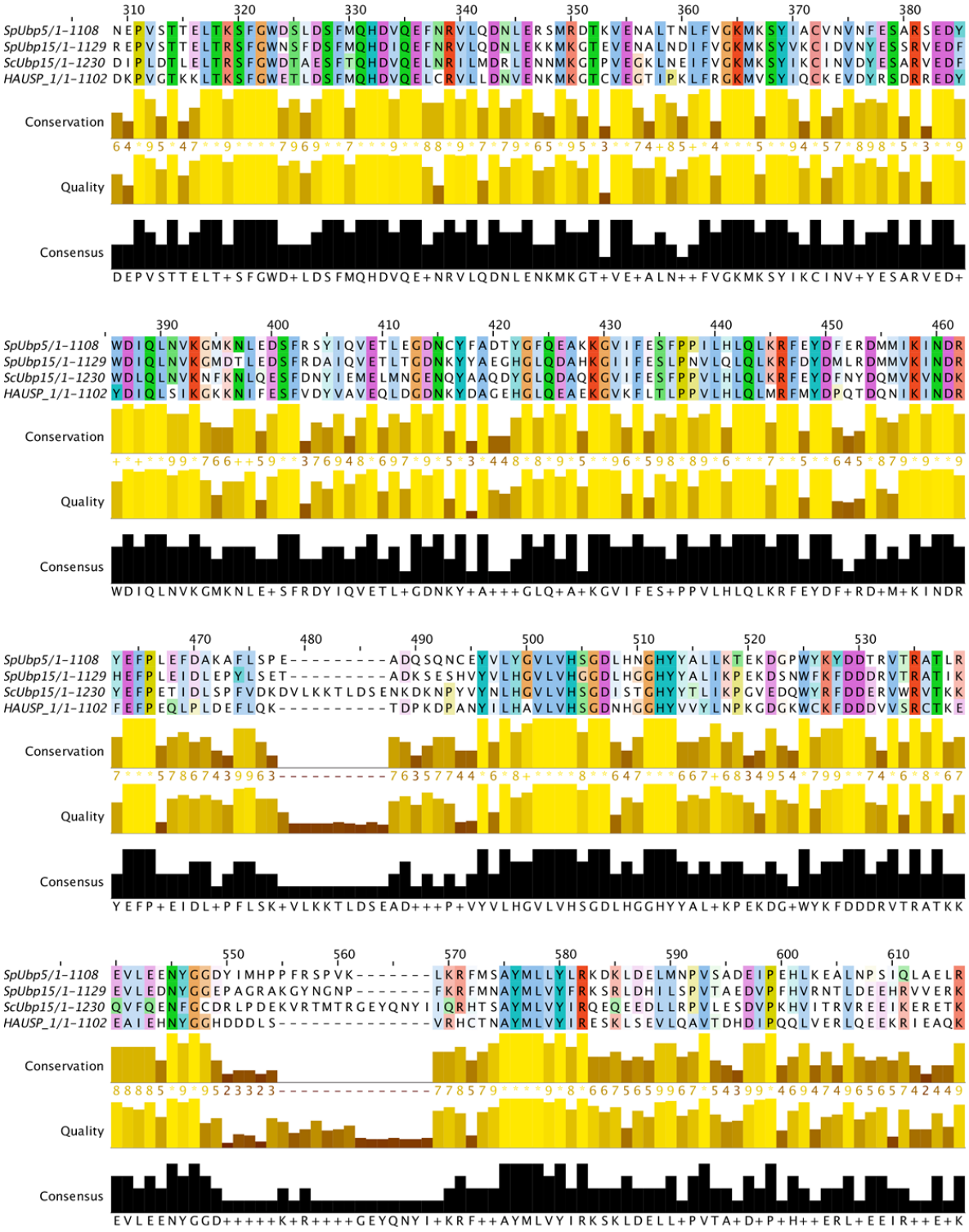
Sde2 co-immunoprecipitates (CoIP) splicing factors. Lysate from *S. pombe Δsde2* strain expressing Sde2–3FLAG epitope-tagged protein was immunoprecipitated using anti-Flag antibody beads. The co-immunoprecipitated proteins were analysed by mass spectrometry (the experiment was repeated twice). The table shows number of unique peptides obtained for each protein in mass spectrometry, and functions of their *S. cerevisiae* orthologs.(The experiment was performed by Shravan Kumar Mishra and Kiran Kumar Kolathur) [200]. Top 75 proteins are shown; nd, not detected; nc, not clear.

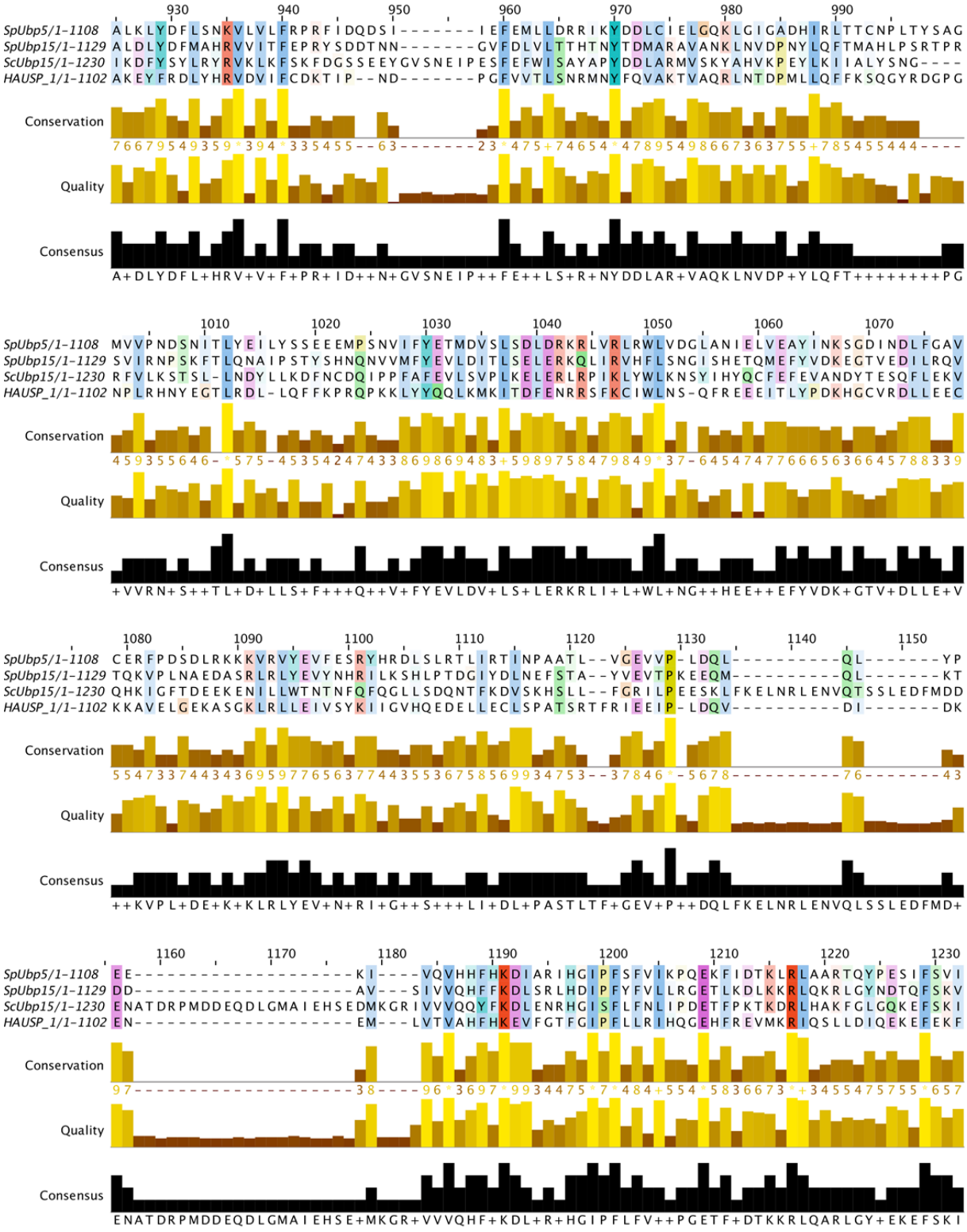
Sr. No.	Protein	No. of unique peptides in Expt. 1	No. of unique peptides in Expt. 2	Putative ortholog in *S. cerevisiae*	Reported in RNA Splicing
1	SPP42	137	48	PRP8	yes
2	CWF10	67	34	SNU114	yes
3	CWF11	66	26	nc	yes
4	BRR2	66	10	Brr2	yes
5	SYF1	55	30	SYF1	yes
6	CDC5	51	23	CEF1	yes
7	PAB1	42	27	PAB1	yes
8	PRP22	42	9	PRP22	yes
9	PRP10	40	7	HSH155	yes
10	CWF4	36	19	CLF1	yes
11	EXO2	36	nd	XRN1	
12	PRP17	32	16	PRP17	yes
13	PRP45	31	10	PRP45	yes
14	PRP12	30	9	RSE1	yes
15	CWF22	29	9	CWC22	yes
16	PRP5	25	1	PRP46	yes
17	SPBC15C4.05	25	5	nc	
18	UCP12	22	nd	YLR419W	
19	PRP19	21	13	PRP19	yes
20	ECM2	19	8	ECM2	yes
21	SPBC4F6.14	19	nd	NOP4	
22	SPAC6G10.07	19	nd	CBC1	yes
23	TCG1	18	10	RNP1	yes
24	CWF17	17	10	nc	yes
25	POL5	17	3	POL5	
26	CWF2	16	10	CWC2	yes
27	SPAC20H4.06C	16	1	nc	yes
28	CWF7	14	13	SNT309	yes
29	SAP145	14	nd	CUS1	yes
30	RPS17	13	6	RPS17	
31	LEA1	13	9	LEA1	yes
32	THI4	13	3	THI6	
33	ERG28	13	2	BUD18	
34	SLU7	12	4	SLU7	yes
35	CWF19	12	nd	nc	yes

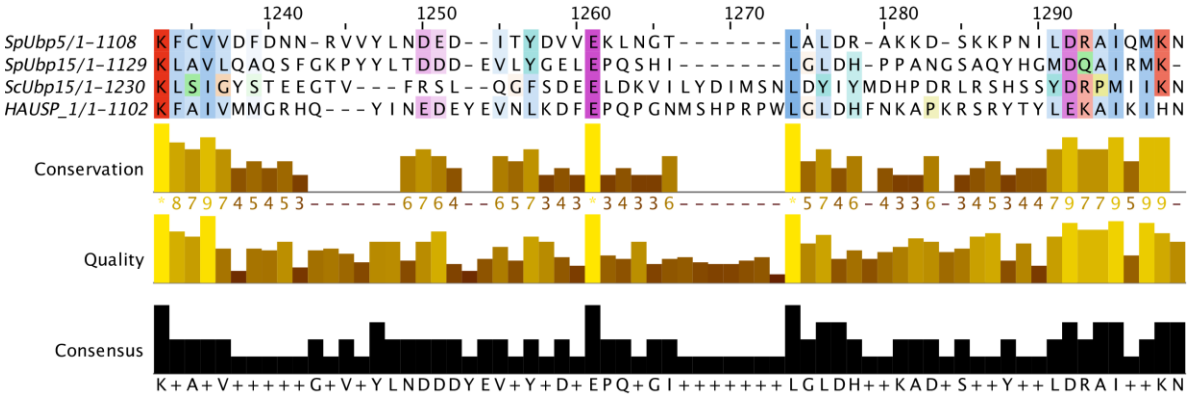
Sr. No.	Protein	No. of unique peptides in Expt. 1	No. of unique peptides in Expt. 2	Putative ortholog in *S. cerevisiae*	Reported in RNA Splicing
36	SPAC20H4.09	12	4	PRP22	yes
37	SPAC1527.03	12	nd	SRO9	
38	RPL26	11	6	RPL26	
39	SMB1	11	6	SMB1	yes
40	SAP61	11	nd	PRP9	yes
41	KIN1	11	nd	KIN1	
42	SYF2	10	7	SYF2	yes
43	CWF12	10	7	ISY1	yes
44	SNU66	10	nd	SNU66	yes
45	CDC28	10	1	PRP2	yes
46	SPAC23C4.17	10	nd	TRM4	
47	SLA1	9	4	LHP1	
48	CWF18	8	3	nc	yes
49	CWF21	8	5	nc	yes
50	RPS24	8	2	RPS24	
51	SMF1	8	1	SMX3	yes
52	CWF24	8	nd	CWC24	yes
53	SMD2	7	6	SMD2	yes
54	CYP1	7	4	nc	
55	THI2	7	3	THI4	
56	SRP2	7	2	NPL3	yes
57	DSK1	7	nd	SKY1	yes
58	NOP12	7	nd	NOP12	
59	PRP1	7	3	PRP6	yes
60	SMD1	6	3	SMD1	yes
61	SMG1	6	3	SMX2	yes
62	MSL1	6	2	MSL1	yes
63	SPAC8F11.04	6	2	CIC1	
64	BAG1	6	3	SNL1	
65	RPL43	6	nd	RPL43	
66	PRP39	6	nd	PRP39	yes
67	SPBP16F5.06	6	nd	NOP8	
68	CWF14	5	3	BUD31	yes
69	SMD3	5	4	SMD3	yes
70	SAP114	5	nd	PRP21	yes
71	RNP24	5	nd	NOP13	
72	SNU71	5	nd	SNU71	yes
73	SDE2	4	3	nc	yes
74	SME1	4	2	SME1	yes
75	CWF15	4	6	CWC15	yes

Figure 5.1 – Multiple sequence alignment of *Sp*Ubp5, *Sp*Ubp15, *Sc*Ubp15 and *Hs*USP7: The alignment was done using Jalview software [203]**.**




SpUbp5/1-1108
SpUbp15/1-1129
ScUbp15/1-1230
HAUSP_1/1-1102
Conservation
Quality
Consensus


930 940 950 960 970 980 990
SpUbp5/1-1108 ALKLYDFLSNKVLVLFRPRFIDQDSI-------IEFEMLLDRRIKYDDLCIELGQKLGIGADHIRLTTCNPLTYSAG
SpUbp15/1-1129 ALDLYDFMAHRVVITFEPRYSDDTNN-------GVFDLVLTTHTNYTDMARAVANKLNVDPNYLQFTMAHLPSRTPR
ScUbp15/1-1230 IKDFYSYLRYRVKLKFSKFDGSSEEYGVSNEIPESFEFWISAYAPYDDLARMVSKYAHVKPEYLKIIALYSNG----
HAUSP_1/1-1102 AKEYFRDLYHRVDVIFCDKTIP--ND-------PGFVVTLSNRMNYFQVAKTVAQRLNTDPMLLQFFKSQGYRDGPG
Conservation
Quality
Consensus
1010 1020 1030 1040 1050 1060 1070
SpUbp5/1-1108 MVVPNDSNITLYEILYSSEEEMPSNVIFYETMDVSLSDLDRKRLVRLRWLVDGLANIELVEAYINKSGDINDLFGAV
SpUbp15/1-1129 SVIRNPSKFTLQNAIPSTYSHNQNVVMFYEVLDITLSELERKQLIRVHFLSNGISHETQMEFYVDKEGTVEDILRQV
ScUbp15/1-1230 RFVLKSTSL-LNDYLLKDFNCDQIPPFAFEVLSVPLKELERLRPIKLYWLKNSYIHYQCFEFEVANDYTESQFLEKV
HAUSP_1/1-1102 NPLRHNYEGTLRDL-LQFFKPRQPKKLYYQQLKMKITDFENRRSFKCIWLNS-QFREEEITLYPDKHGCVRDLLEEC
Conservation
Quality
Consensus
1080 1090 1100 1110 1120 1130 1140 1150
SpUbp5/1-1108 CERFPDSDLRKKKVRVYEVFESRYHRDLSLRTLIRTINPAATL--VGEVVP-LDQL----------QL-------YP
SpUbp15/1-1129 TQKVPLNAEDASRLRLYEVYNHRILKSHLPTDGIYDLNEFSTA--YVEVTPKEEQM----------QL-------KT
ScUbp15/1-1230 QHKIGFTDEEKENILLWTNTNFQFQGLLSDQNTFKDVSKHSLL--FGRILPEESKLFKELNRLENVQTSSLEDFMDD
HAUSP_1/1-1102 KKAVELGEKASGKLRLLEIVSYKIIGVHQEDELLECLSPATSRTFRIEEIP-LDQV----------DI-------DK
Conservation
Quality
Consensus
1160 1170 1180 1190 1200 1210 1220 1230
SpUbp5/1-1108 EE----------KI----VQVHHFHKDIARIHGIPFSFVIKPQEKFIDTKLRLAARTQYPESIFSVI
SpUbp15/1-1129 DD----------AV---SIVVQHFFKDLSRLHDIPFYFVLLRGETLKDLKKRLQKRLGYNDTQFSKV
ScUbp15/1-1230 ENATDRPMDDEQDLGMAIEHSEDMKGRIVVVQQYFKDLENRHGISFLFNLIPDETFPKTKDRLHAKFGLGQKEFSKI
HAUSP_1/1-1102 EN----------EM---LVTVAHFHKEVFGTFGIPFLLRIHQGEHFREVMKRIQSLLDIQEKEFEKF
Conservation
Quality
Consensus



➢ **Figure related to section 3.11 - Ubiquitin like processing of Sde2 is conserved in humans**

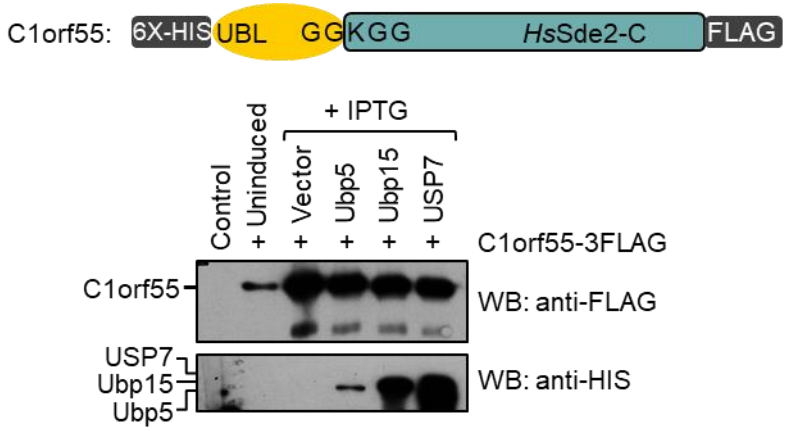


Figure 3.11.2 – Co-expression of C1orf55 (*Hs*Sde2) with *S. pombe* Ubp5, Ubp15 and *Hs*USP7
Expression constructs harbouring indicated gene or cDNAs were co-transformed in *E. coli* BL21 (DE3) strain. Following protein expression, total cell lysates were processed by immunoblotting using anti-FLAG for detection of *Hs*Sde2 and anti-His antibody for detection of proteases.

References

1 Clague, M.J. and Urbé, S. (2010) Ubiquitin: Same molecule, different degradation pathways. *Cell* 143, 682–685

2 Ciechanover, A. and Schwartz, A. L. (1998) The ubiquitin-proteasome pathway: the complexity and myriad functions of proteins death. *Proc. Natl. Acad. Sci. U. S. A.* 95, 2727–2730

3 Lacombe, T. *et al.* (2009) Linear ubiquitin fusion to Rps31 and its subsequent cleavage are required for the efficient production and functional integrity of 40S ribosomal subunits. *Mol. Microbiol.* 72, 69–84

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

5 Hershko, A. (1998) The ubiquitin system. In: Peters JM., Harris J.R., Finley D. (eds) *Ubiquitin Biol. Cell*. Springer, Boston, MA

6 Finley, D. *et al.* (1989) The tails of ubiquitin precursors are ribosomal proteins whose fusion to ubiquitin facilitates ribosome biogenesis. *Nature* 338, 394–401

7 Xu, P. *et al.* (2009) Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* 137, 133–145

8 Peng, J. *et al.* (2003) A proteomics approach to understanding protein ubiquitination. TL - 21. *Nat. Biotechnol.* 21, 921–926

9 Shang, F. *et al.* (2005) Lys6-modified ubiquitin inhibits ubiquitin-dependent protein degradation. *J. Biol. Chem.* 280(21), 20365-20374

10 Vucic, D. *et al.* (2011) Ubiquitylation in apoptosis: A post-translational modification at the edge of life and death. *Nature Reviews Molecular Cell Biology* 12(7), 439-452

11 Livneh, I. *et al.* (2016) The life cycle of the 26S proteasome: From birth, through regulation and function, and onto its death. *Cell Res.* 26, 869–885

12 Díaz-Villanueva, J.F. *et al.* (2015) Protein folding and mechanisms of proteostasis. *Int. J. Mol. Sci.* 16, 17193–17230

13 Groll, M. *et al.* Structure of 20S proteasome from yeast at 2.4 Å resolution. , *Nature*, 386. (1997) , 463–471

14 Dick, T.P. *et al.* (1998) Contribution of proteasomal?-subunits to the cleavage of peptide substrates analyzed with yeast mutants. *J. Biol. Chem.* 273, 25637–25646

15 Li, X. and Demartino, G.N. (2009) Variably modulated gating of the 26S proteasome by ATP and polyubiquitin. *Biochem. J.* 421(3), 397–404

16 Kleijnen, M.F. *et al.* (2007) Stability of the proteasome can be regulated allosterically through engagement of its proteolytic active sites. *Nat. Struct. Mol. Biol.* 14, 1180–1188

17 Clague, M.J. *et al.* (2013) Deubiquitylases From Genes to Organism. *Physiol. Rev.* 93, 1289–1315

18 Reyes-Turcu, F.E. *et al.* (2009) Regulation and Cellular Roles of Ubiquitin-Specific Deubiquitinating Enzymes. *Annu. Rev. Biochem.* 78, 363–397

19 Mevissen, T.E.T. and Komander, D. (2017) Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* 86, 159–192

20 Abdul Rehman, S.A. *et al.* (2016) MINDY-1 Is a Member of an Evolutionarily Conserved and Structurally Distinct New Family of Deubiquitinating Enzymes. *Mol. Cell* 63, 146–155

21 Hickey, C.M. *et al.* (2012) Function and regulation of SUMO proteases. *Nat. Rev. Mol. Cell Biol.* 13, 755–766

22 Shin, E.J. *et al.* (2012) DeSUMOylating isopeptidase: A second class of SUMO protease. *EMBO Rep.* 13, 339–346

23 Nijman, S.M.B. *et al.* (2005) A genomic and functional inventory of deubiquitinating enzymes. *Cell* 123, 773–786

24 Hu, M. *et al.* (2005) Structure and mechanisms of the proteasome-associated deubiquitinating enzyme USP14. *EMBO J.* 24, 3747–3756

25 Hu, M. *et al.* (2002) Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* 111, 1041–1054

26 Ye, Y. *et al.* (2009) Dissection of USP catalytic domains reveals five common insertion points.

Mol. Biosyst. 5, 1797–1808
27 Komander, D. (2010) Mechanism, Specificity and Structure of the Deubiquitinases. *Conjug. Deconjugation Ubiquitin Fam. Modif.* 54, 69-87
28 Setsuie, R. and Wada, K. (2007) The functions of UCH-L1 and its relation to Neurodegenerative diseases. *Neurochemistry International*. 51(2-4), 105-111
29 Leroy, E. *et al.* (1998) The ubiquitin pathway in Parkinson's disease. *Nature*. 395(6701),451-452
30 Yao, T. *et al.* (2006) Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1. *Nat. Cell Biol.* 8, 994–1002
31 Hamazaki, J. *et al.* (2006) A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes. *EMBO J.* 25(19), 4524-4536
32 Mallery, D.L. *et al.* (2002) Activation of the E3 ligase function of the BRCA1/BARD1 complex by polyubiquitin chains. *EMBO J.* 21, 6755–6762
33 Nishikawa, H. *et al.* (2009) BRCA1-associated protein 1 interferes with BRCA1/BARD1 RING heterodimer activity. *Cancer Res.* 69, 111–119
34 Komander, D. *et al.* (2009) Breaking the chains: Structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* 10, 550–563
35 Finley, D. (2009) Recognition and Processing of Ubiquitin-Protein Conjugates by the Proteasome. *Annu. Rev. Biochem.* 78, 477-513
36 Williams, R.L. and Urbé, S. (2007) The emerging shape of the ESCRT machinery. *Nat. Rev. Mol. Cell Biol.* 8, 355
37 Clague, M.J. and Urbé, S. (2006) Endocytosis: the DUB version. *Trends Cell Biol.* 16, 551–559
38 Cope, G.A. *et al.* (2002) Role of predicted metalloprotease motif of Jab1/Csn5 in cleavage of Nedd8 from Cul1. *Science.* 298, 608–611
39 Pena, V. *et al.* (2007) Structure of a Multipartite Protein-Protein Interaction Domain in Splicing Factor Prp8 and Its Link to Retinitis Pigmentosa. *Mol. Cell* 25, 615–624
40 Iyer, L.M. *et al.* (2004) Novel predicted peptidases with a potential role in the ubiquitin signaling pathway. *Cell Cycle* 3(11), 1440-1150
41 Sheedlo, M.J. *et al.* (2015) Structural basis of substrate recognition by a bacterial deubiquitinase important for dynamics of phagosome ubiquitination. *Proc. Natl. Acad. Sci.* 112, 15090–15095
42 Sato, Y. *et al.* (2015) Structures of CYLD USP with Met1-or Lys63-linked diubiquitin reveal mechanisms for dual specificity. *Nat. Struct. Mol. Biol.* 22, 222–229
43 Ronau, J.A. *et al.* (2016) Substrate specificity of the ubiquitin and Ubl proteases. *Cell Res.* 26, 441–456
44 Morgan, M.T. *et al.* (2016) Structural basis for histone H2B deubiquitination by the SAGA DUB module. *Science.* 351, 725–728
45 Sahtoe, D.D. and Sixma, T.K. (2015) Layers of DUB regulation. *Trends Biochem. Sci.* 40, 456–467
46 Sokol V. Todi, and H.L.P. (2011) Balancing Act: Deubiquitinating Enzymes in the Nervous System. *Changes* 34, 370–382
47 Lopez-Castejon, G. and Edelmann, M.J. (2016) Deubiquitinases: Novel Therapeutic Targets in Immune Surveillance. *Mediators Inflamm.* 2016, 3481371
48 Pinto-Fernandez, A. and Kessler, B.M. (2016) DUBbing cancer: Deubiquitylating enzymes involved in epigenetics, DNA damage and the cell cycle as therapeutic targets. *Front. Genet.* 7, 1–13
49 Coornaert, B. *et al.* (2008) T cell antigen receptor stimulation induces MALT1 paracaspase - Mediated cleavage of the NF-κB inhibitor A20. *Nat. Immunol.* 9, 263–271
50 Huang, T.T. *et al.* (2006) Regulation of monoubiquitinated PCNA by DUB autocleavage. *Nat. Cell Biol.* 8, 339–347
51 Urbe, S. *et al.* (2012) Systematic survey of deubiquitinase localization identifies USP21 as a regulator of centrosome- and microtubule-associated functions. *Mol. Biol. Cell* 23, 1095–1103
52 Kouranti, I. *et al.* (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 Biol.* 8,
53 Mizuno, E. *et al.* (2007) 14-3-3-dependent inhibition of the deubiquitinating activity of UBPY and its cancellation in the M phase. *Exp. Cell Res.* 313, 3624–3634
54 Zhang, X. and Wang, Y. (2015) Cell cycle regulation of VCIP135 deubiquitinase activity and function in p97/p47-mediated Golgi reassembly. *Mol. Biol. Cell* 26, 2242–2251
55 Demuc, A. *et al.* (2009) The UBA-UIM domains of the USP25 regulate the enzyme ubiquitination state and modulate substrate recognition. *PLoS One* 4,
56 Lee, J.G. *et al.* (2013) Reversible inactivation of deubiquitinases by reactive oxygen species in vitro and in cells. *Nat. Commun.* 4, 1512–1568
57 Kulathu, Y. *et al.* (2013) Regulation of A20 and other OTU deubiquitinases by reversible oxidation. *Nat. Commun.* 4, 1569
58 Cotto-Rios, X.M. *et al.* (2012) Deubiquitinases as a Signaling Target of Oxidative Stress. *Cell Rep.* 2, 1475–1484
59 van der Veen, A.G. and Ploegh, H.L. (2012) Ubiquitin-like proteins. *Annu Rev Biochem* 81, 323–357
60 Rabut, G. *et al.* (2008) Function and regulation of protein neddylation. "Protein modifications: beyond the usual suspects" review series. *EMBO Rep.* 118, 83–97
61 Watson, I.R. *et al.* (2011) NEDD8 Pathways in Cancer, Sine Quibus Non. *Cancer Cell* 19, 168–176
62 Xirodimas, D.P. *et al.* (2004) Mdm2-mediated NEDD8 conjugation of p53 inhibits its transcriptional activity. *Cell* 118, 83–97
63 Perng, Y.C. and Lenschow, D.J. (2018) ISG15 in antiviral immunity and beyond. *Nat. Rev. Microbiol.* 16, 423–439
64 Jeon, Y.J. *et al.* (2010) ISG15 and immune diseases. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1802, 485–496
65 Mathers, C. *et al.* (2014) The Human Cytomegalovirus UL26 Protein Antagonizes NF- B Activation. *J. Virol.* 88, 14289–14300
66 Kim, Y.J. *et al.* (2016) Consecutive Inhibition of ISG15 Expression and ISGylation by Cytomegalovirus Regulators. *PLoS Pathog.* 12, 1–28
67 Rahnefeld, A. *et al.* (2014) Ubiquitin-like protein ISG15 (Interferon-Stimulated Gene of 15 kDa) in host defense against heart failure in a mouse model of virus-induced cardiomyopathy. *Circulation* 130, 1589–1600
68 Villarroya-Beltri, C. *et al.* (2016) ISGylation controls exosome secretion by promoting lysosomal degradation of MVB proteins. *Nat. Commun.* 7,
69 Okumura, A. *et al.* (2006) Innate antiviral response targets HIV-1 release by the induction of ubiquitin-like protein ISG15. *Proc. Natl. Acad. Sci.* 103, 1440–1445
70 Matunis, M.J. *et al.* (1996) A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *J. Cell Biol.* 135, 1457–1470
71 Mahajan, R. *et al.* (1998) Molecular characterization of the SUMO-1 modification of RanGAP1 and its role in nuclear envelope association. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *J Cell Biol* 140, 259–70.
72 Müller, S. *et al.* (2001) Sumo, ubiquitin's mysterious cousin. *Nat. Rev. Mol. Cell Biol.* 2, 202–210
73 Bayer, P. *et al.* (1998) Structure determination of the small ubiquitin-related modifier SUMO-1. *J. Mol. Biol.* 280, 275–286
74 Yeh, E.T.H. *et al.* (2000) Ubiquitin-like proteins: New wines in new bottles. *Gene* 248, 1–14
75 Chakrabarti, S.R. *et al.* (2000) Posttranslational modification of TEL and TEL/AML1 by SUMO-1 and cell-cycle-dependent assembly into nuclear bodies. *Proc. Natl. Acad. Sci. U. S. A.* 97, 13281–13285
76 Müller, S. *et al.* (2000) POST-TRANSLATION MODIFICATION AND DEGRADATION : c-Jun and p53 Activity Is Modulated by SUMO-1 Modification c-Jun and p53 Activity Is Modulated by SUMO-1 Modification. 275, 13321–13329
77 Kim, Y.H. *et al.* (1999) Covalent modification of the homeodomain-interacting protein kinase

2 (HIPK2) by the ubiquitin-like protein SUMO-1. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12350–5
78 Fogal, V. *et al.* (2000) Regulation of p53 activity in nuclear bodies by a specific PML isoform. *EMBO J.* 19, 6185–95
79 Desterro, J.M. *et al.* (1998) SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol. Cell* 2, 233–239
80 Buschmann, T. *et al.* (2000) SUMO-1 modification of Mdm2 prevents its self-ubiquitination and increases Mdm2 ability to ubiquitinate p53. *Cell* 101, 753–762
81 Schmidtke, G. *et al.* (2014) FAT10ylation as a signal for proteasomal degradation. *Biochim. Biophys. Acta - Mol. Cell Res.* 1843, 97–102
82 Basler, M. *et al.* (2015) The ubiquitin-like modifier FAT10 in antigen processing and antimicrobial defense. *Mol. Immunol.* 68, 129–132
83 Aichem, A. and Groettrup, M. (2016) The ubiquitin-like modifier FAT10 in cancer development. *Int. J. Biochem. Cell Biol.* 79, 451–461
84 Spinnenhirn, V. *et al.* (2014) The ubiquitin-like modifier FAT10 decorates autophagy-targeted Salmonella and contributes to Salmonella resistance in mice. *J. Cell Sci.* 127, 4883–4893
85 Geng, J. and Klionsky, D.J. (2008) The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. "Protein Modifications: Beyond the Usual Suspects" Review Series. *EMBO Rep.* 9, 859–864
86 Jüdes, A. *et al.* (2015) Urmylation and tRNA thiolation functions of ubiquitin-like Uba4·Urm1 systems are conserved from yeast to man. *FEBS Lett.* 589, 904–909
87 Pedrioli, P.G.A. *et al.* (2008) ' Protein Modifications : Beyond the Usual Suspects ' Review Series. *EMBO Rep.* 9, 1196–1202
88 Wang, F. *et al.* (2011) The dual role of ubiquitin-like protein Urm1 as a protein modifier and sulfur carrier. *Protein Cell* 2, 612–619
89 Hochstrasser, M. (2009) Origin and function of ubiquitin-like proteins. *Nature* 458, 422–429
90 Komatsu, M. *et al.* (2004) A novel protein-conjugating system for Ufm1, a ubiquitin-fold modifier. *EMBO J.* 23, 1977–1986
91 Sung, H.K. *et al.* (2007) Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J. Biol. Chem.* 282, 5256–5262
92 Wei, Y. and Xu, X. (2016) UFMylation: A Unique & Fashionable Modification for Life. *Genomics, Proteomics Bioinforma.* 14, 140–146
93 Pick, E. *et al.* (2009) PCI Complexes: Beyond the Proteasome, CSN, and eIF3 Troika. *Mol. Cell* 35, 260–264
94 Yoo, H.M. *et al.* (2014) Modification of asc1 by ufm1 is crucial for erα transactivation and breast cancer development. *Mol. Cell* 56, 261–274
95 Daniel, J. and Liebau, E. (2014) The Ufm1 Cascade. *Cells* 3, 627–638
96 Wilkinson, C.R.M. *et al.* (2004) Ubiquitin-like protein Hub1 is required for pre-mRNA splicing and localization of an essential splicing factor in fission yeast. *Curr. Biol.* 14(24), 2283-2288
97 Yashiroda, H. and Tanaka, K. (2004) Hub1 is an essential ubiquitin-like protein without functioning as a typical modifier in fission yeast. *Genes to Cells* 9, 1189–1197
98 Mishra, S.K. *et al.* (2011) Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. *Nature* 474, 173–180
99 Karaduman, R. *et al.* (2017) Error-Prone Splicing Controlled by the Ubiquitin Relative Hub1. *Mol. Cell* 67, 423–432
100 Benedetti, C. *et al.* (2006) Ubiquitin-Like Protein 5 Positively Regulates Chaperone Gene Expression in the Mitochondrial Unfolded Protein Response. 239, 229–239
101 Ammon, T. *et al.* (2014) The conserved ubiquitin-like protein Hub 1 plays a critical role in splicing in human cells. *J. Mol. Cell Biol.* 6, 312–323
102 Oka, Y. *et al.* (2015) Ubiquitin-like protein UBL5 promotes the functional integrity of the Fanconi anemia pathway. *EMBO J.* 34(10), 1385-1398
103 Matera, A.G. and Wang, Z. (2014) A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* 15, 108–121
104 Will, C.L. and Lührmann, R. (2011) Spliceosome structure and function. TL-3. *Cold Spring Harb. Perspect. Biol.* 3, 1–23

105 Wahl, M.C. *et al.* (2009) The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* 136, 701–718
106 Patel, A.A. and Steitz, J.A. (2003) Splicing double: Insights from the second spliceosome. *Nat. Rev. Mol. Cell Biol.* 4, 960–970
107 Fox-Walsh, K.L. *et al.* (2005) The architecture of pre-mRNAs affects mechanisms of splice-site pairing. *Proc. Natl. Acad. Sci.* 102, 16176–16181
108 Berglund, J.A. *et al.* (1998) A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes Dev.* 12, 858–867
109 Zamore, P.D. *et al.* (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature* 355, 609–614
110 Graveley, B.R. *et al.* (2001) The role of U2AF 35 and U2AF 65 in enhancer-dependent splicing.
111 Bessonov, S. *et al.* (2010) Characterization of purified human B. *Rna* 16, 2384–2403
112 Sun, J.-S. and Manley, J.L. (1995) A novel {U2}-{U6} {snRNA} structure is necessary for mammalian {mRNA} splicing. *Genes Dev.* 9, 843–854
113 Henriques, J.A.P. and Moustacchi, E. (1980) Isolation and characterization of pso mutants sensitive to photo-addition of psoralen derivatives in Saccharomyces cerevisiae. *Genetics.* 95(2), 273-288
114 Chanarat, S. and Sträßer, K. (2013) Splicing and beyond: The many faces of the Prp19 complex. *Biochim. Biophys. Acta - Mol. Cell Res.* 1833, 2126–2134
115 Ohi, M.D. *et al.* (2003) Structural insights into the U-box, a domain associated with multi-ubiquitination. *Nat. Struct. Biol.* 10, 250–255
116 Song, E.J. *et al.* (2010) The Prp19 complex and the Usp4Sart3deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev.* 24, 1434–1447
117 Sheth, N. *et al.* (2006) Comprehensive splice-site analysis using comparative genomics. *Nucleic Acids Res.* 34, 3955–3967
118 Roca, X. *et al.* (2005) Determinants of the inherent strength of human 5′ splice sites. *Rna* 11, 683–698
119 Zhang, X.H.F. *et al.* (2005) Dichotomous splicing signals in exon flanks. *Genome Res.* 15, 768-779
120 Belkadi, A. *et al.* (2013) The Incidentalome. *Nature* 8, 1–8
121 Gelfman, S. *et al.* (2012) Changes in exon-intron structure during vertebrate evolution affect the splicing pattern of exons. *Genome Res.* 22, 35–50
122 Berget, S.M. (1995) Exon recognition in vertebrate splicing. *Journal of Biological Chemistry*. 270(6), 2411-2414
123 Hoffman, B.E. and Grabowski, P.J. (1992) U1 snRNP targets an essential splicing factor, U2AF65 to the 3' splice sit by a network of interactions spanning the exon. *Genes Dev.* 6(12B), 2554-68
124 Reed, R. Mechanisms of fidelity in pre-mRNA splicing. , *Current Opinion in Cell Biology*. (2000)
125 De Conti, L. *et al.* (2013) Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip. Rev. RNA* 4, 49–60
126 Keren, H. *et al.* (2010) Alternative splicing and evolution: Diversification, exon definition and function. *Nat. Rev. Genet.* 11, 345–355
127 Wang, E.T. *et al.* (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476
128 Tazi, J. *et al.* (2009) Alternative splicing and disease. *Biochim. Biophys. Acta - Mol. Basis Dis.* 1792, 14–26
129 Wang, G.S. and Cooper, T.A. (2007) Splicing in disease: Disruption of the splicing code and the decoding machinery. *Nat. Rev. Genet.* 8, 749–761
130 Barbosa-Morais, N.L. *et al.* (2012) The evolutionary landscape of alternative splicing in vertebrate species. *Science* 338(6114), 1587-1593
131 Lee, Y. and Rio, D.C. (2015) Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu. Rev. Biochem.* 84, 291–323
132 Fu, X.D. and Ares, M. (2014) Context-dependent control of alternative splicing by RNA-

binding proteins. *Nat. Rev. Genet.* 15, 689–701
133 Chen, M. and Manley, J.L. (2009) Mechanisms of alternative splicing regulation: Insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell Biol.* 10, 741–754
134 Mishra, S.K. and Thakran, P. (2018) Intron specificity in pre-mRNA splicing. *Curr. Genet.* 64, 777–784
135 Kim, E. *et al.* (2008) Alternative splicing: Current perspectives. *BioEssays* 30, 38–47
136 Koren, E. *et al.* (2007) The Emergence of Alternative 3′ and 5′ Splice Site Exons from Constitutive Exons. *PLoS Comput. Biol.* 3, e95
137 Änkö, M.L. (2014) Regulation of gene expression programmes by serine-arginine rich splicing factors. *Seminars in Cell and Developmental Biology*. 32, 11-21
138 Zhou, Z. and Fu, X.D. (2013) Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma.* 122(3), 191-207
139 Martinez-Contreras, R. *et al.* (2007) hnRNP proteins and splicing control. *Adv. Exp. Med. Biol.* 623, 123–147
140 Cartegni, L. *et al.* (2002) Listening to silence and understanding nonsense: Exonic mutations that affect splicing. *Nature Reviews Genetics*. 3(4), 285-98
141 Zhu, J. *et al.* (2001) Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol. Cell* 8(6), 1351-1361.
142 Blanchette, M. (1999) Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J.* 18(7), 1939–1952
143 Heiner, M. *et al.* (2010) HnRNP L-mediated regulation of mammalian alternative splicing by interference with splice site recognition. *RNA Biol.* 7(1), 56-64
144 Förch, P. *et al.* (2002) The splicing regulator TIA-1 interacts with U1-C to promote U1 snRNP recruitment to 5′ splice sites. *EMBO J.* 21, 6882–6892
145 Buratti, E. and Baralle, F.E. (2004) Influence of RNA Secondary Structure on the Pre-mRNA Splicing Process. *Mol. Cell. Biol.* 24(24), 10505-10514
146 Rouskin, S. *et al.* (2014) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. *Nature* 505(7485), 701-705
147 Graveley, B.R. (2005) Mutually exclusive splicing of the insect Dscam Pre-mRNA directed by competing intronic RNA secondary structures. *Cell* 123(1), 65-73
148 May, G.E. *et al.* (2011) Competing RNA secondary structures are required for mutually exclusive splicing of the Dscam exon 6 cluster. *RNA RNA* 17(2):222-229
149 Kishore, S. *et al.* (2006) Regulation of alternative splicing by snoRNAs. *Cold Spring Harb. Symp. Quant. Biol.* 71,329-334
150 Kishore, S. and Stamm, S. (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* 311(5758), 230-232
151 Sims, R.J. *et al.* (2007) Recognition of Trimethylated Histone H3 Lysine 4 Facilitates the Recruitment of Transcription Postinitiation Factors and Pre-mRNA Splicing. *Mol. Cell* 28, 665–676
152 Beswick, J.A. and Rizzo, C. (2008) Regulation of Alternative Splicing by Histone Modifications. *Ann. la Fond. Louis Broglie* 33, 31–43
153 Sowa, M.E. *et al.* (2009) Defining the Human Deubiquitinating Enzyme Interaction Landscape. *Cell* 138, 389–403
154 Barboric, M. *et al.* (2009) 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc. Natl. Acad. Sci.* 106, 7798 -7803
155 Trippe, R. *et al.* (2006) Identification, cloning, and functional analysis of the human U6 snRNA-specific terminal uridylyl transferase. *RNA* 12(8), 1494-1504
156 Bellare, P. *et al.* (2008) A role for ubiquitin in the spliceosome assembly pathway. *Nat. Struct. Mol. Biol.* 15, 444–451
157 Das, T. *et al.* (2017) USP15 regulates dynamic protein-protein interactions of the spliceosome through deubiquitination of PRP31. *Nucleic Acids Res.* 45, 4866–4880
158 Singh, R.K. and Cooper, T.A. (2012) Pre-mRNA splicing in disease and therapeutics. *Trends in Molecular Medicine*. 18(8), 472–482.

159 Padgett, R.A. (2012) New connections between splicing and human disease. *Trends in Genetics*. 28(4), 147-54
160 Boon, K.L. *et al.* (2007) Prp8 mutations that cause human retinitis pigmentosa lead to a U5 snRNP maturation defect in yeast. *Nat. Struct. Mol. Biol.* 14(11), 1077-1083
161 Tanackovic, G. *et al.* (2011) PRPF mutations are associated with generalized defects in spliceosome formation and pre-mRNA splicing in patients with retinitis pigmentosa. *Hum. Mol. Genet.* 20(11), 2116-30
162 Utz, V.M. *et al.* (2013) Autosomal Dominant Retinitis Pigmentosa Secondary to Pre-mRNA Splicing-Factor Gene PRPF31 (RP11): Review of Disease Mechanism and Report of a Family with a Novel 3-Base Pair Insertion. *Ophthalmic Genet.* 34, 183–188
163 Schrank, B. *et al.* (1997) Inactivation of the survival motor neuron gene, a candidate gene for human spinal muscular atrophy, leads to massive cell death in early mouse embryos. *Proc. Natl. Acad. Sci. U. S. A.* 94(18), 9920-9925
164 Lorson, C.L. *et al.* (1999) A single nucleotide in the SMN gene regulates splicing and is responsible for spinal muscular atrophy. *Proc. Natl. Acad. Sci.* 96(11), 6307-6311
165 Bäumer, D. *et al.* (2009) Alternative splicing events are a late feature of pathology in a mouse model of spinal muscular atrophy. *PLoS Genet.* 5(12), e1000773
166 Cazzola, M. *et al.* (2013) Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms. *Blood*. 121(2), 260-269
167 Yoshida, K. *et al.* (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 78(7367), 64-69
168 Chesnais, V. *et al.* (2012) Spliceosome mutations in myelodysplastic syndromes and chronic myelomonocytic leukemia. *Oncotarget* 3(11):1284-1293
169 He, H. *et al.* (2011) Mutations in U4atac snRNA, a component of the minor spliceosome, in the developmental disorder MOPD i. *Science* 332(6026), 238-240
170 Makarov, E.M. *et al.* (2002) Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science* 298(5601), 2205-2208
171 Rappsilber, J. *et al.* (2003) Large-scale proteomic analysis of the human spliceosome. *Genome Res.* 13, 1231–1245
172 Tagwerker, C. *et al.* (2006) A Tandem Affinity Tag for Two-step Purification under Fully Denaturing Conditions. *Mol. Cell. Proteomics* 5, 737–748
173 Bellare, P. *et al.* (2006) Ubiquitin binding by a variant Jab1/MPN domain in the essential pre-mRNA splicing factor Prp8p. *Rna* 12, 292–302
174 Hatakeyama, S. *et al.* (2001) U Box Proteins as a New Family of Ubiquitin-Protein Ligases. *J. Biol. Chem.* 276, 33111–33120
175 Pozzi, B. *et al.* (2017) SUMO conjugation to spliceosomal proteins is required for efficient pre-mRNA splicing. *Nucleic Acids Res.* 45, 6729–6745
176 Pelisch, F. *et al.* (2010) The serine/arginine-rich protein SF2/ASF regulates protein sumoylation. *Proc. Natl. Acad. Sci.* 107, 16119–16124
177 Blomster, H.A. *et al.* (2009) Novel Proteomics Strategy Brings Insight into the Prevalence of SUMO-2 Target Sites. *Mol. Cell. Proteomics* 8, 1382–1390
178 Vertegaal, A.C.O. *et al.* (2004) A proteomic study of SUMO-2 target proteins. *J. Biol. Chem.* 279, 33791–33798
179 Kim, D.-U. *et al.* (2010) Analysis of a genome-wide set of gene deletions in the fission yeast Schizosaccharomyces pombe. *Nat Biotech.* 28, 617–623
180 Kennedy, P.J. *et al.* (2008) A genome-wide screen of genes involved in cadmium tolerance in Schizosaccharomyces pombe. *Toxicol. Sci.* 106, 124–139
181 Zhang, L. *et al.* (2013) Genome-wide screening for genes associated with valproic acid sensitivity in fission yeast. *PLoS One* 8, e68738
182 Sugioka-Sugiyama, R. and Sugiyama, T. (2011) Sde2: A novel nuclear protein essential for telomeric silencing and genomic stability in Schizosaccharomyces pombe. *Biochem. Biophys. Res. Commun.* 406, 444–8
183 Tadeo, X. *et al.* (2013) Elimination of shelterin components bypasses RNAi for pericentric heterochromatin assembly. *Genes Dev.* 27, 2489–2499
184 Bayne, E.H. *et al.* (2014) A systematic genetic screen identifies new factors influencing

centromeric heterochromatin integrity in fission yeast. *Genome Biol.* 15, 481
185 Yang, J. *et al.* (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Meth* 12, 7–8
186 Bessonov, S. *et al.* (2008) Isolation of an active step I spliceosome and composition of its RNP core. *Nature* 452, 846–850
187 Chen, W. *et al.* (2014) Endogenous U2·U5·U6 snRNA complexes in S. pombe are intron lariat spliceosomes. *RNA* 20, 308–20
188 Jo, U. *et al.* (2016) PCNA-Dependent Cleavage and Degradation of SDE2 Regulates Response to Replication Stress. *PLoS Genet.* 12, 1–26
189 Knop, M. *et al.* (1999) Epitope tagging of yeast genes using a PCR-based strategy: More tags and improved practical routines. *Yeast* 15, 963–972
190 Janke, C. *et al.* (2004) A versatile toolbox for PCR-based tagging of yeast genes: New fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* 21, 947–962
191 Vjestica, A. *et al.* (2016) Microscopy of Fission Yeast Sexual Lifecycle. *J. Vis. Exp.* 109, e53801
192 Inada, M. and Pleiss, J.A. (2010) Genome-wide approaches to monitor pre-mrna splicing. *Methods Enzymol.* 470, 51–75
193 Chen, W. *et al.* (2014) Endogenous U2•U5•U6 snRNA complexes in S. pombe are intron lariat spliceosomes. *Rna* 20, 308–320
194 Nguyen Ba, A.N. *et al.* (2009) NLStradamus: A simple Hidden Markov Model for nuclear localization signal prediction. *BMC Bioinformatics* 10, 1–11
195 Matsuyama, A. *et al.* (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. *Nat. Biotechnol.* 24, 841–847
196 Cheng, Y. *et al.* (2016) Characterization of a Ran gene from Puccinia striiformis f. sp. tritici involved in fungal growth and anti-cell death. *Sci. Rep.* 6, 1–11
197 Jallepalli *et al.* (1998) sud1(+) targets cyclin-dependent kinase-phosphorylated Cdc18 and Rum1 proteins for degradation and stops unwanted diploidization in fission yeast. *Proc. Natl. Acad. Sci. U. S. A.* 95, 8159–8164
198 Richert, K. *et al.* (2002) The deubiquitinating enzyme Ubp21p of fission yeast stabilizes a mutant form of protein kinase Prp4p. *Mol. Genet. Genomics* 267, 88–95
199 Debelyy, M.O. *et al.* (2011) Ubp15p, a ubiquitin hydrolase associated with the peroxisomal export machinery. *J. Biol. Chem.* 286, 28223–28234
200 Thakran, P. *et al.* (2017) Sde2 is an intron-specific pre-mRNA splicing regulator activated by ubiquitin-like processing. *EMBO J.* 37, 89-101
201 Varshavsky, A. (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci.* 20, 1298–1345
202 Brooks, C.L. *et al.* (2007) The p53-Mdm2-HAUSP complex is involved in p53 stabilization by HAUSP. *Oncogene* 26(51):7262-7366
203 Waterhouse, A.M. *et al.* (2009) Jalview Version 2-A multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191
204 Kim, R.Q. *et al.* (2016) Structure of USP7 catalytic domain and three Ubl-domains reveals a connector α-helix with regulatory role. *J. Struct. Biol.* 195, 11–18
205 Nurse, P. *et al.* (1976) Genetic control of the cell division cycle in the fission yeast Schizosaccharomyces pombe. *MGG Mol. Gen. Genet.* 146, 167–178
206 Burns, C.G. *et al.* (2002) Evidence that Myb-related CDC5 proteins are required for pre-mRNA splicing. *Proc. Natl. Acad. Sci.* 96, 13789–13794
207 Collier, S.E. *et al.* (2014) Structural and functional insights into the N-terminus of Schizosaccharomyces pombe Cdc5. *Biochemistry* 53, 6439–6451
208 Hogg, R. *et al.* (2010) The function of the NineTeen Complex (NTC) in regulating spliceosome conformations and fidelity during pre-mRNA splicing. *Biochem. Soc. Trans.* 38, 1110–1115
209 Lorenzi, L.E. *et al.* (2015) Fission yeast Cactin restricts telomere transcription and elongation by controlling Rap1 levels. *EMBO J.* 34, 115–129
210 Zanini, I.M.Y. *et al.* (2017) Human cactin interacts with DHX8 and SRRM2 to assure efficient pre-mRNA splicing and sister chromatid cohesion. *J. Cell Sci.* 130, 767-778

211 Doherty, M.F. *et al.* (2014) Proteomic Analysis Reveals CACN-1 Is a Component of the Spliceosome in *Caenorhabditis elegans*. *G3: Genes|Genomes|Genetics* 4, 1555–1564
212 Baltz, A.G. *et al.* (2012) The mRNA-Bound Proteome and Its Global Occupancy Profile on Protein-Coding Transcripts. *Mol. Cell* 46, 674–690
213 Beckley, J.R. *et al.* (2015) A Degenerate Cohort of Yeast Membrane Trafficking DUBs Mediates Cell Polarity and Survival. *Mol. Cell. Proteomics* 14, 3132–3141
214 Li, M. *et al.* (2002) Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* 416(6881), 648-53
215 Cummins, J.M. and Vogelstein, B. (2004) HAUSP is required for p53 destabilization. *Cell Cycle* 3, 689–692
216 Hu, M. *et al.* (2006) Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: Implications for the regulation of the p53-MDM2 pathway. *PLoS Biol.* 4(2), e27
217 Everett, R.D. *et al.* (1997) A novel ubiquitin-specific protease is dynamically associated with the PML nuclear domain and binds to a herpesvirus regulatory protein. *EMBO J.* 16(7), 1519–1530
218 Zapata, J.M. *et al.* (2007) CHAPTER 1 Phylogeny of the TRAF / MATH Domain.
219 Sheng, Y. *et al.* (2006) Molecular recognition of p53 and MDM2 by USP7/HAUSP. *Nat. Struct. Mol. Biol.* 13, 285–291
220 Saridakis, V. *et al.* (2005) Structure of the p53 binding domain of HAUSP/USP7 bound to epstein-barr nuclear antigen 1: Implications for EBV-mediated immortalization. *Mol. Cell* 18, 25–36
221 Bozza, W.P. and Zhuang, Z. (2011) Biochemical characterization of a multidomain deubiquitinating enzyme Ubp15 and the regulatory role of its terminal domains. *Biochemistry* 50, 6423–6432
222 Schauber, C. *et al.* (1998) Rad23 links DNA repair to the ubiquitin/proteasome pathway. *Nature* 391, 715–718
223 Hartmann-Petersen, R. *et al.* (2003) Transferring substrates to the 26S proteasome. *Trends Biochem. Sci.* 28, 26–31
224 Madsen, L. *et al.* (2007) Ubiquitin domain proteins in disease. *BMC Biochem.* 8, 1–8
225 Matheny, S.A. *et al.* (2004) Ras regulates assembly of mitogenic signalling complexes through the effector protein IMP. *Nature* 427, 256–260
226 Hayes, S.D. *et al.* (2012) Direct and indirect control of mitogen-activated protein kinase pathway-associated components, BRAP/IMP E3 ubiquitin ligase and CRAF/RAF1 kinase, by the deubiquitylating enzyme USP15. *J. Biol. Chem.* 287, 43007–43018
227 Clerici, M. *et al.* (2014) The DUSP-Ubl domain of USP4 enhances its catalytic efficiency by promoting ubiquitin exchange. *Nat. Commun.* 5, 1–11
228 Foster, B.M. *et al.* (2018) Critical Role of the UBL Domain in Stimulating the E3 Ubiquitin Ligase Activity of UHRF1 toward Chromatin. *Mol. Cell* 72, 739–752
229 Rougé, L. *et al.* (2016) Molecular Understanding of USP7 Substrate Recognition and C-Terminal Activation. *Structure* 24, 1335–1345
230 Kim, R.Q. *et al.* (2016) Mechanism of USP7/HAUSP activation by its C-Terminal ubiquitin-like domain and allosteric regulation by GMP-synthetase. *Mol. Cell* 281, 147–159
231 Komander, D. *et al.* (2008) The Structure of the CYLD USP Domain Explains Its Specificity for Lys63-Linked Polyubiquitin and Reveals a B Box Module. *Mol. Cell* 29, 451–464
232 Catic, A. *et al.* (2007) Screen for ISG15-crossreactive deubiquitinases. *PLoS One* 2,
233 Basters, A. *et al.* (2017) Structural basis of the specificity of USP18 toward ISG15. *Nat. Struct. Mol. Biol.* 24, 270–278
234 Ye, Y. *et al.* (2011) Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. *EMBO Rep.* 12, 350–357
235 Schulz, S. *et al.* (2012) Ubiquitin-specific protease-like 1 (USPL1) is a SUMO isopeptidase with essential, non-catalytic functions. *EMBO Rep.* 13, 930–938
236 Pruneda, J.N. *et al.* (2016) The Molecular Basis for Ubiquitin and Ubiquitin-like Specificities in Bacterial Effector Proteases. *Mol. Cell* 63, 261–276

237 Love, K.R. *et al.* (2007) Mechanisms, biology and inhibitors of deubiquitinating enzymes. *Nat. Chem. Biol.* 3, 697–705
238 Konarska, M.M. and Query, C.C. (2005) Insights into the mechanisms of splicing: More lessons from the ribosome. *Genes and Development*. 19:2255-2260
239 Staley, J.P. and Woolford, J.L. (2009) Assembly of ribosomes and spliceosomes: complex ribonucleoprotein machines. *Curr. Opin. Cell Biol.* 21, 109–118

Publication

Co-first authored publication -

- Thakran P, **Pandit PA**, Datta S, Kolathur KK, Pleiss JA, and Mishra SK (2017). Sde2 is an intron-specific pre-mRNA splicing regulator activated by ubiquitin-like processing. ***EMBO J***. 37, 89-101