Charles University

Faculty of Science

Study programme: Special Chemical and Biological Programmes

Branch of study: Molecular Biology and Biochemistry of Organisms

Tereza Hubáčková

Role of the ubiquitin-like protein, Hub1, in the pre-mRNA splicing regulation

Role proteinu podobného ubiquitinu, Hub1, v regulaci sestřihu pre-mRNA

Bachelor´s thesis

Supervisor: Mgr. Kateřina Abrhámová, Ph.D.

Prague, 2018

Prohlášení

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

V Praze, 10. 5. 2018

Tereza Hubáčková

Acknowledgements

I would like to express sincere gratitude to my supervisor, Mgr. Kateřina Abrhámová, Ph.D., for her valuable advices and support, and to my family for endless support and inspiration.

ABSTRACT

Splicing is a key step of eukaryotic gene expression and as well as other steps of this vital process, splicing has to be tightly regulated. Hub1 protein is a ubiquitin-like protein which noncovalently interacts with spliceosomal proteins Snu66 and Prp5 ATPase. According to the proposed model, low level of Hub1 protein stimulates ATPase activity of Prp5 helicase sufficiently for splicing of optimal splice sites, but not for splicing of suboptimal ones. Nevertheless, high level of Hub1 protein stimulates Prp5 ATPase sufficiently for splicing of both splice-site types. Excessive level of Hub1 protein may be harmful for the cell, because the immoderate splicing of suboptimal splice sites may produce aberrantly-spliced transcripts as a by-product. Hub1-induced negative feedback loop safeguards the cell from Hub1 protein hyperactivity by regulation of Prp5 ATPase level. Additionally, Hub1 protein regulates alternative splicing of *Saccharomyces cerevisiae SRC1* gene and ensures appropriate balance of its products.

KEY WORDS

Hub1 protein, splicing regulation, UBL5 protein, ubiquitin-like protein (UBL), Prp5 helicase, Snu66 protein, *Saccharomyces cerevisiae*

ABSTRAKT

Sestřih je klíčovým krokem eukaryotické genové exprese, a stejně tak jako ostatní kroky tohoto životně důležitého procesu, musí být pevně regulován. Protein Hub1 je proteinem podobným ubiquitinu, který nekovalentně interaguje se spliceosomovými proteiny Snu66 a Prp5 ATPázou. Podle navrženého modelu nízká hladina proteinu Hub1 stimuluje ATPázovou aktivitu Prp5 proteinu dostatečně pro sestřih optimálních sestřihových míst, ale už ne pro sestřih suboptimálních míst. Nicméně vysoká hladina Hub1 proteinu je dostatečná pro stimulaci Prp5 ATPázy pro sestřih obou typů sestřihových míst. Jelikož sestřih suboptimálních substrátů může generovat jako vedlejší produkt aberantně sestřižené transkripty, přílišná aktivita proteinu Hub1 by mohla být pro buňku škodlivá. Před rizikem hyperaktivního proteinu Hub1 chrání buňku negativní zpětnovazebná smyčka regulující hladinu Prp5 ATPázy. Tato smyčka je proteinem Hub1 indukována. Navíc Hub1 protein reguluje alternativní sestřih genu *SRC1 Saccharomyces cerevisiae* a zajišťuje tak náležitou rovnováhu jeho produktů.

KLÍČOVÁ SLOVA

protein Hub1, regulace sestřihu, protein UBL5, protein podobný ubiquitinu, Prp5 helikáza, protein Snu66, *Saccharomyces cerevisiae*

LIST OF ABBREVIATIONS

AA(s)	amino acid(s)
ADP	adenosine diphosphate
ATP	adenosine triphosphate
ATPase	ATP hydrolyzing enzyme
BP	branchpoint
cDNA	coding DNA
Cryo-EM	cryo-electron microscopy
C-terminus	carboxy terminus of a protein
DEAD-box	family of ATPases carrying characteristic motif of AAs (aspartate, glutamate, alanine, aspartate)
DExD/H-box	family of ATPases carrying characteristic motif of AAs (aspartate, glutamate, any AA, aspartate or histidine)
DNA	deoxyribonucleic acid
HIND	Hub1-interaction domain
mRNA	messenger RNA
N-terminus	amino terminus of a protein
ORF	open reading frame
pre-mRNA	precursor messenger RNA
RMSD	root-mean-square deviation
RNA(s)	ribonucleic acid(s)
RNP(s)	ribonucleoprotein(s)
RT-qPCR	quantitative polymerase chain reaction with reverse transcription
SIM	SUMO interaction motif
snRNA(s)	small nuclear RNA(s)
snRNP	small nuclear ribonucleoprotein
Sulfo-MBS	m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester
UBL(s)	ubiquitin-like protein(s)
UIM	ubiquitin interaction motif
wt	wild type
3D	three dimensional
3’SS	3’splice site
5’SS	5’splice site

TABLE OF CONTENTS

1 INTRODUCTION

Transition of genetic information from deoxyribonucleic acid (DNA) to protein is mediated by ribonucleic acids (RNAs). At first, DNA is transcribed into precursor messenger RNA (pre-mRNA). This process takes place in the nucleus as far as eukaryotic organisms are considered. Before the pre-mRNA is allowed to leave the nucleus, it is processed: typically capped, polyadenylated and often also spliced. Finally, the mature messenger RNA (mRNA) emerges.

Splicing is the cellular process by which noncoding introns (intervening sequences) are removed. Interestingly, the splicing of particular intron may happen using different splice sites and this state is called alternative splicing. This process allows emergence of many distinct protein products from a single transcript, but on the other hand the complexity of the whole process provides many opportunities for disease-causing errors, for example in case of Huntington disease (Sathasivam et al., 2013). Our understanding of splicing and particularly its regulation not only broadens our knowledge of nature but also helps to treat diseases in which splicing dysregulation is the primary cause.

The aim of my bachelor thesis is to review what is known about the role of an unconventional ubiquitin-like protein Hub1 in splicing regulation.

2 THE COMPLEX PROCESS OF SPLICING

Splicing in eukaryotes is mostly performed by spliceosome. This colossal machinery is a protein-directed metalloribozyme which is composed of five small nuclear RNAs (snRNAs), specifically U1, U2, U4/U6 duplex and U5, which are in complexes with many ribonucleoproteins (RNPs). Additional myriad of non-ribonucleoproteins and Mg^{2+} ions are present. Splicing reaction consists of two transesterifications, the 1st step is known as branching and the 2nd step is called exon ligation (Figure 1a). Spliceosomal complex undergoes multiple complicated structural rearrangements between at least nine stages (Figure 1b) in order to execute these two reactions. Critical step of this process is the recognition of authentic pre-mRNA splicing signals: 5'splice site (5'SS), 3'splice site (3'SS) and branchpoint sequence (BP sequence, only in Figure 1 it is designated exceptionally as BS) (*Shi, 2017).

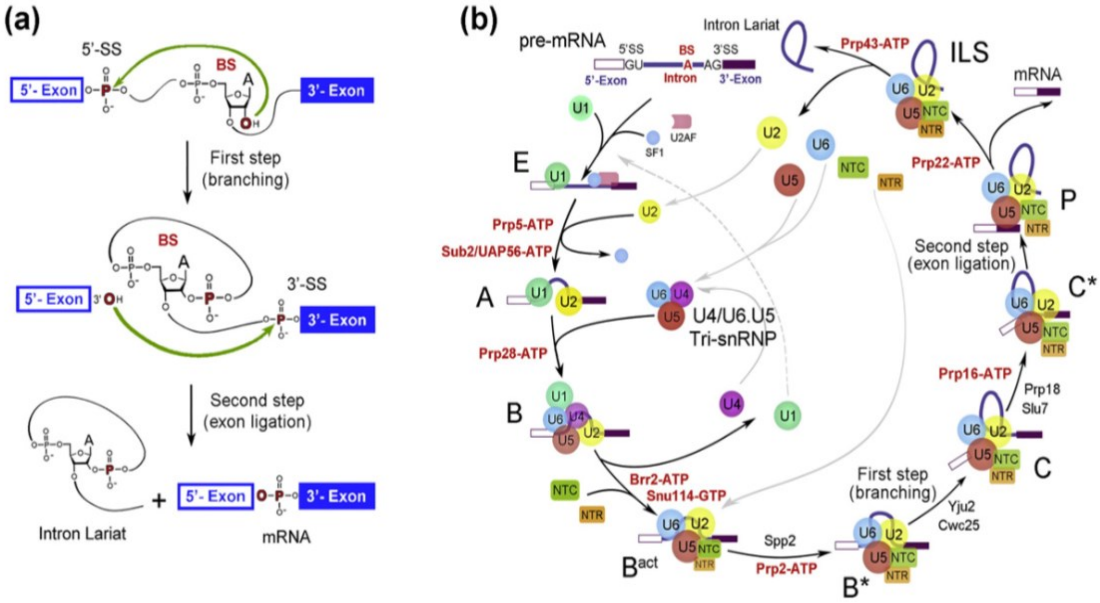

Figure 1: Two-step splicing reaction (a) and spliceosomal cycle (b). (a) Branchpoint adenine initiates nucleophilic attack on the 5'phosphate of terminal intronic guanine and subsequently the 3'hydroxyl of the last ribonucleotide of 5'exon attacks 5'phosphate of the 3'exon. (b) Spliceosomal cycle shows dynamics of various components between spliceosomal states. Names and corresponding abbreviations of the complexes are as follows: E – early spliceosome, A – pre-spliceosome, B – pre-catalytic complex, B^{act} – activated complex, B* - catalytically activated complex, C – the catalytic step I spliceosome, C* - the step II catalytically activated complex, P – post-catalytic complex, ILS – intron lariat spliceosome. Figure taken from (*Shi, 2017).

At the very beginning of a spliceosomal cycle in *Saccharomyces cerevisiae* (*S. cerevisiae*), 5'end invariant ACUUAC hexaribonucleotide of U1 snRNA contacts through base pairing 5'SS of pre-mRNA (*Newman, 2008). 5'splice sites of distinct introns differ in sequences with the most usual sequence in *S. cerevisiae* being GUAUGU. This sequence is called consensus and it is, except for its 4th position, complementary to the invariant region of U1 (*Chiou and Lynch, 2014).

At the same time, or possibly even before attachment of U1 small nuclear ribonucleoprotein (snRNP), branchpoint sequence interacts with heterodimer of Mud2 and BBP (branchpoint binding protein) and this interaction is soon replaced by U2 snRNA in resulting pre-spliceosome complex A (*Herzel et al., 2017). Invariant region of U2 snRNA GUAGUA forms a duplex with BP UACUAAC consensus sequence. The mismatch between the branchpoint adenosine (designated in red) and the 1st guanosine of U2 snRNA (given in blue) bulges out the branchpoint adenosine into favorable position for nucleophilic attack. However, before the branching reaction is allowed to happen, several structural rearrangements proceed (*Newman, 2008).

Tri-snRNP, consisting of U4/U6 duplex with freely associated U5, joins the spliceosome which causes transition into B complex. Subsequently about 30 proteins are recruited to pre-catalytic complex B and U6 snRNA replaces U1 snRNA in its binding with 5'SS. This new contact is mediated by the last three ribonucleotides of ACAGAG invariant element of U6 snRNA (*Newman, 2008). The U5 snRNA directly binds 5' exon by its GCCUUUUAC (*Newman, 2008) and thus helps to hold pre-mRNA in proper position for both future catalytic steps. U1 and U4 snRNAs leave the B complex and activated complex B^{act} emerges (*Shi, 2017).

Spliceosome cycle continues in similar fashion with multiple rearrangements until the mature mRNA is freed and components of spliceosome together with intron-lariat RNA are processed and possibly recycled (*Shi, 2017), but from the perspective of this work, particularly the splice site recognition and the early events described above are of major importance. Between 2015 and 2018 many spliceosomal structures captured at different stages at atomic or near atomic resolution were published. These structures promise to provide deep molecular-level insight into the mechanisms of structural rearrangements and function of single parts of this enormous complex. Cryo-electron microscopy (cryo-EM) has already allowed for detailed view of pre-catalytic complex B of human and budding yeast (Bertram et al., 2017; Plaschka et al., 2017), activated complex B^{act} (Yan et al., 2016; Zhang et al., 2018), structure of spliceosome immediately after branching reaction (Galej et al., 2016), the catalytic step I spliceosome C (Wan et al., 2016), also the step II catalytically activated complex C* (Yan et al., 2017), post-catalytic complex P (Bai et al., 2017) and intron lariat spliceosome ILS (Wan et al., 2017). However, the detailed structures of E and A complexes of comparably high resolution are missing.

Key components of spliceosome machinery not yet mentioned are DExD/H-box ATPases/helicases. Eight spliceosomal ATPases direct dynamic rearrangements of spliceosome by promoting both, creation and disruption of RNA-RNA or RNA-protein interactions (*Semlow and Staley, 2012). Repeated inspecting of splicing signals in early stages of the spliceosomal cycle is critical for recognition of right splicing substrates. 5'SS interacts firstly with U1 snRNA, but it is later doublechecked by base pairing with U6 snRNA. BP sequence is at first recognized by Mud2/BBP heterodimer but then again controlled by interaction with U2 snRNP (*Herzel et al., 2017). In principle, spliceosomal helicases participate on the selection process by a specific type of kinetic proofreading. To explain the concept in brief, an alternative non-productive branch in the spliceosomal pathway exists beside the productive catalytic pathway. RNA helicases may either favor this alternative pathway for certain substrate, leading to rejection or even discard of this substrate, or favor the continuation of the productive splicing pathway. This activity of helicases may be regulated globally or locally to relax the specificity and allow the use of suboptimal nonconsensus splice sites (*Semlow and Staley, 2012).

For this work, specifically Prp5 helicase is particularly important. It is a DEAD-box helicase, conserved from *S. cerevisiae* to *H. sapiens*, which is engaged in early spliceosome assembly (Figure 1b). Most importantly, Prp5p inspects binding of U2 snRNA to BP sequence and either promotes an alternative pathway if the interaction is suboptimal or promotes the productive step, transition into A complex. In this way, Prp5 enzyme ensures quality control of BP sequence and contributes to splicing fidelity (*Semlow and Staley, 2012) It was suggested that dissociation of Prp5p after the stable interaction between U2 snRNA and BP sequence probably allow association of tri-snRNP (Liang and Cheng, 2015).

The second protein of major importance from the perspective of this work is Snu66p. Both of these proteins, Snu66p and Prp5p, interact with our protein of interest, Hub1 (Karaduman et al., 2017; Mishra et al., 2011). Snu66p is one of the components of tri-snRNP (Nguyen et al., 2016; Stevens and Abelson, 1999) which is a preassembled complex that binds to the pre-spliceosome, induces transition into the pre-catalytic complex and thus finishes the process of spliceosome assembly (*Shi, 2017). Apart from participation in pre-mRNA splicing (Stevens et al., 2001) Snu66p is involved in processing of 5S ribosomal RNA (Li et al., 2009).

3 FROM SEQUENCE TO MECHANISM OF INTERACTION

Sequence of *S. cerevisiae* chromosome XIV was published in 1997 (Philippsen et al., 1997) and among many other open reading frames (ORFs), *YNR032C-A* was identified on this chromosome. In 2001, this ORF was determined as one of the homologs of a newly characterized human protein Ubiquitin-like 5 (UBL5). UBL5 was discovered when a screen of human iris coding DNA (cDNA) library for a potential causative protein of ocular disease was performed. It was supposed that proteins both highly expressed and conserved in different species in the same tissue would have profound tissue-specific significance and might cause pathologies. Specifically, the porcine choroidal ring cDNA probe was used. UBL5 was given its name on the basis of predicted ubiquitin-like fold. (Friedman et al., 2001). Later this prediction was confirmed, UBL5 possesses ubiquitin fold and it is therefore a member of the ubiquitin-like protein family (UBLs) (McNally et al., 2003). UBL5 was identified in human pre-catalytic complex B (Deckert et al., 2006), but neither UBL5 nor the *S. cerevisiae* homolog is displayed in cryo-EM structures of B-complex (Bertram et al., 2017; Plaschka et al., 2017).

3.1 Sequence comparison

3.1.1 Interspecies comparison of homologs

Product of *YNR032C-A* of *S. cerevisiae* is termed Hub1 protein (homologous to ubiquitin 1) (Dittmar et al., 2002) and its homologs in mammals are known as Beacon (Collier et al., 2000) or alternatively as UBL5 (Friedman et al., 2001). Hub1 is relatively small protein, 8 270.6 Da in size, 73 amino acids (AAs) long, with isoelectric point 6.51 (Saccharomyces genome database, accessed 2018). Whereas in *S. cerevisiae* Hub1p is nonessential (Dittmar et al., 2002), homologs in *S. pombe* and human cells are essential (Ammon et al., 2014; Wilkinson et al., 2004; Yashiroda and Tanaka, 2004).

Hub1 protein is remarkably conserved between eukaryotes, which is demonstrated by the protein sequence alignment from 10 distinct eukaryotic species, see Figure 2 (Karaduman et al., 2017). Sequence identity of Hub1 homologs when compared to human UBL5 ranges between 64.4 % in *S. cerevisiae* (Friedman et al., 2001) over 74 % in *Schizosaccharomyces pombe* (Wilkinson et al., 2004) to 80.8 % in *Caenorhabditis elegans* (Friedman et al., 2001). High level of conservation throughout the phylogeny suggests important role of Hub1 protein (Wilkinson et al., 2004).

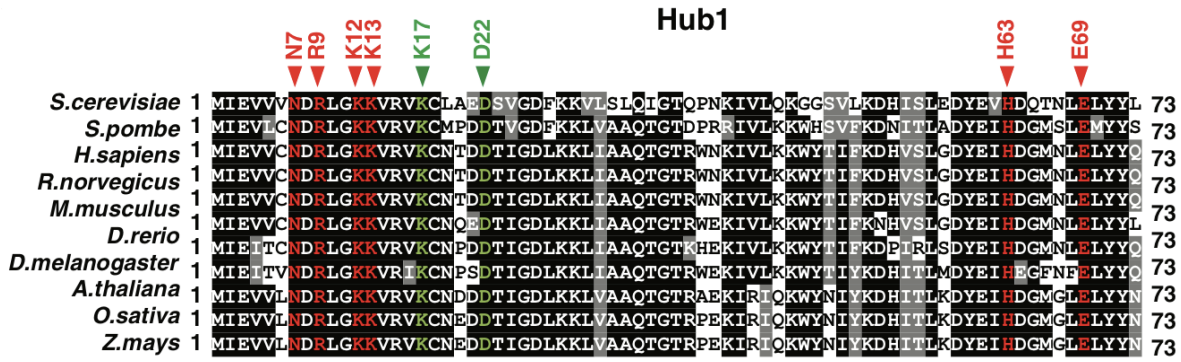

Figure 2: Sequence alignment of Hub1 homologs in ten eukaryotic species. Alignment of Hub1p homologs shows conserved dityrosine motif (positions 71 and 72). Green arrowheads designate conserved residues of Surface I, which are critical for interaction with HIND (Hub1 interaction domain). Aspartate residue (D22) residue is of particular importance. Red arrowheads mark conserved residues of Surface II, which are important for binding to Prp5p. Histidine-63 (H63) is important for this interaction. Note the carboxy-terminal (C-terminal) nonconserved residue. Darker background corresponds to higher conservation. Figure taken from (Karaduman et al., 2017).

3.1.2 Comparison with ubiquitin

Hub1 protein was given its name because of the predicted high structural similarity to ubiquitin, however, its sequence identity to ubiquitin is low (Dittmar et al., 2002). In contrast to the high sequence conservation between Hub1p homologs, Hub1p displays only slight sequence identity to ubiquitin, given in numbers 22 %, which is close to the limit of statistical significance (Dittmar et al., 2002).

Striking feature that results from the sequence comparison with ubiquitin is the lack of conserved diglycine motif in all compared Hub1 homologs (Dittmar et al., 2002; Karaduman et al., 2017). This motif is otherwise present in ubiquitin and some other UBLs, for example Smt3/SUMO-1 or Rub1/NEDD8 and it mediates the covalent conjugation of UBLs to their protein targets (human homologs are designated after slash) (*Hochstrasser, 2009). Unique feature of Hub1 homologs is possession of conserved dityrosine motif near their carboxy terminus (C-terminus), see Figure 2, (Lüders et al., 2003).

3.2 3D structure

The first three dimensional (3D) structures of Hub1 in *S. cerevisiae* and *H. sapiens* were determined in 2003 (McNally et al., 2003; Ramelot et al., 2003) and the ubiquitin fold previously predicted (Friedman et al., 2001) was thus confirmed. Typical β-grasp fold consists of a half open β-barrel with two flanking α-helices. This fold is also described as $\beta\beta\alpha\beta\alpha\beta$ pattern.

Ubiquitin-like fold of both human and budding-yeast was confirmed later by determination at higher resolution. (Ammon et al., 2014; Mishra et al., 2011). Quantified similarity between ubiquitin and Hub1 protein was computed for main chain heavy atoms as RMSD 0.88 Å (root-mean-square deviation) (Mishra et al., 2011).

3.2.1 Comparison with ubiquitin and other UBLs

Pronounced difference between Hub1p and ubiquitin is a protruding C-terminus. Whereas C-terminus of ubiquitin is outstretched from the compact protein core, that of Hub1p is relatively hidden (Figure 3). Comparison with other UBLs, Rub1 and SUMO-1 or its yeast homolog Smt3, gives similar results. In contrast to Hub1p, C-terminuses of all these proteins are also unstructured, relatively freely moving around compact protein core (Lüders et al., 2003).

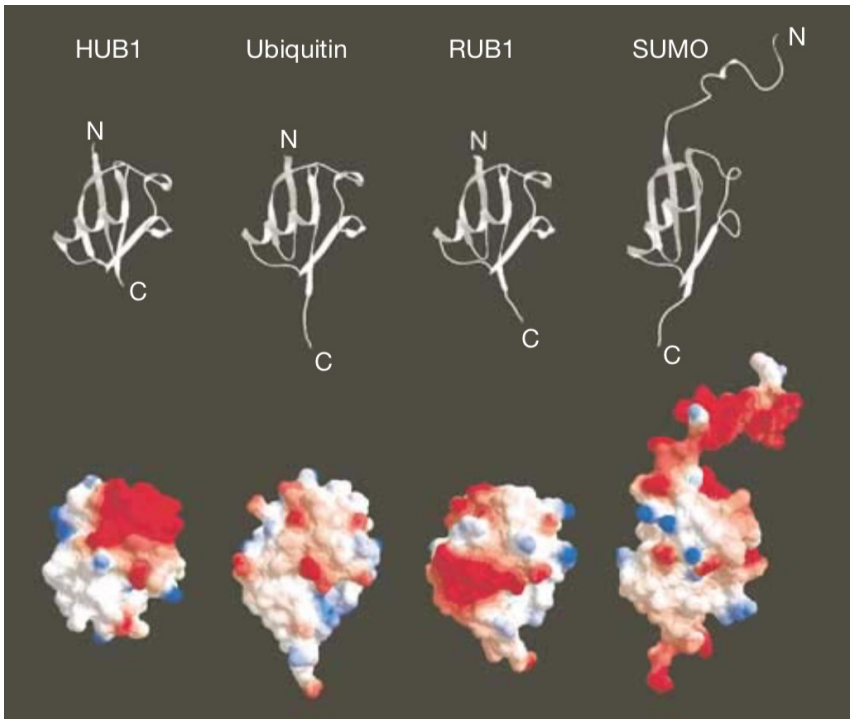

Figure 3: Structural comparison of ubiquitin and ubiquitin-like proteins. Hub1p lacks protruding C-terminus, in contrast to Rub1, SUMO and ubiquitin. Notes: Hub1 of *S. cerevisiae* (Ramelot et al., 2003), Rub1/NEDD8 of *H. sapiens* (Whitby et al., 1998), ubiquitin of *H. sapiens* (Vijay-Kumar et al., 1987), SUMO-1 of *H. sapiens* (Bayer et al., 1998). Figure taken from (Lüders et al., 2003).

Functional consequences might be deduced from this structural difference. Ubiquitin and other UBLs use C-terminus for conjugation to target proteins (*Hochstrasser, 2009). C-terminus of UBL5 (or Hub1p) wouldn't be easily available for this conjugation, which suggests the possibility of another interaction mechanism (McNally et al., 2003).

3.3 Mechanism of interaction with other proteins

Major structural similarity between ubiquitin and Hub1 protein led to initial concept of Hub1p as a covalent protein modifier (Dittmar et al., 2002), another one in a row of UBLs that covalently modify protein targets via C-terminus (*Hochstrasser, 2009). The main question was, how exactly the modification is done, when Hub1p lacks conserved diglycine motif and when Hub1p's very last C-terminal residue is not conserved at all. However, the conserved dityrosine motif (residues 71 and 72) near the C-terminus seemed to be a plausible candidate for conjugation (Dittmar et al., 2002).

3.3.1 The first model

After series of experiments in *S. cerevisiae*, a model of covalent conjugation through tyrosine-72 after the proteolytic cleavage of nonconserved residue-73 was proposed. Thus tyrosine-72 was considered to be critical for function of Hub1p (Dittmar et al., 2002). However, another research group followed the published protocols of Dittmer et al. and repeated these experiments. They claim that they could not confirm those results, but the data were not published (Lüders et al., 2003).

Nevertheless, plenty of results in contradiction with this initial concept of covalent conjugation were published later. Hub1p was reported to create SDS-resistant adducts and this process was not abolished by concurrent terminal tagging on both terminuses of Hub1 protein. Both tags were detected in adducts, which means that even if natural C-terminus was not available to mediate conjugation, adducts still formed. This suggests different mechanism of conjugation. Moreover, the generation of these adducts was not ATP-dependent (Lüders et al., 2003). Most importantly, this ATP-independence argues against a ubiquitin-like conjugation mechanism, because this mechanism is essentially ATP-dependent (*Hochstrasser, 2009). C-tagged Hub1p, but not free Hub1p, was detected also in recombinant *S. pombe*. This observation also opposes the model of C-terminal conjugation. (Yashiroda and Tanaka, 2004).

Dispensability of dityrosine motif, which was initially considered crucial because of its conservation throughout phylogeny, was repeatedly verified in *S. pombe*, where Hub1 protein is essential. Truncated Hub1p lacking both tyrosines was able to rescue lethality, which means that its essential function does not require those tyrosines (Wilkinson et al., 2004; Yashiroda and Tanaka, 2004).

3.3.2 The second model

A new concept of modification performed by Hub1p was suggested. The model of noncovalent binding to protein targets (Lüders et al., 2003). This idea was verified later both in *S. cerevisiae* (Karaduman et al., 2017; Mishra et al., 2011) and human cell lines (Ammon et al., 2014).

Similarly as for covalent conjugation model, it was demonstrated, that also some known noncovalent interactions of Hub1p are independent of conserved dityrosine motif (Mishra et al., 2011). However, distinct C-terminal modifications do affect functionality of Hub1 protein.

Removal of four AAs or exchange of dityrosine motif for two lysine residues, but not for other AAs, prevent successful complementation of essential function in *S. pombe* (Wilkinson et al., 2004). Artificial extension of two aspartic acid residues (Hub1-DD) fails in the same way. Moreover, this Hub1-DD modified variant is functionally defective in *S. cerevisiae*. Resulting phenotype is for example the loss of ability to provide splicing of certain transcripts or alternative splicing of *SRC1* gene (Mishra et al., 2011) (discussed in detail in chapter 5.1.3).

The reason why the C-terminal extension of only two residues disrupts the function of Hub1 protein is not currently clear on molecular level, but the easiest explanation seems to be that overall structural stability of Hub1p is affected.

4 DETAILS OF NONCOVALENT INTERACTION BETWEEN HUB1 AND ITS TWO PROTEIN PARTNERS

Hub1p possibly interacts with myriad of other proteins. However, for a splicing-related function of Hub1p, two interactors are currently known to be particularly important. They are Snu66p and Prp5p (Karaduman et al., 2017; Mishra et al., 2011).

4.1 Interaction with Snu66 protein

The first clue to the interaction of Hub1p and Snu66p was the phenotypic rescue of temperature sensitive allele *hub1-1* by overproduction of Snu66p in *S. pombe* (Yashiroda and Tanaka, 2004). Interaction between these two proteins was later detected by two-hybrid screen in *S. cerevisiae* (Wilkinson et al., 2004) and two-hybrid assay (Mishra et al., 2011).

Short N-terminal segment of Snu66p is responsible for Hub1p binding. This element is termed HIND which stands for Hub1-interaction domain. Isolated HIND adopts helical conformation in solution, which is otherwise unusual for short fragments. Hub1-HIND interface covers approximately 500 $Å^2$ and most of the interactions are hydrophobic contacts, although one salt bridge is present (Mishra et al., 2011).

Snu66 protein of *S. cerevisiae* carries two highly similar HIND elements in tandem (72 % sequence identity). These are called HIND-I, which is 18 AAs long, and HIND-II which is 19 AAs long. Surprisingly, Snu66 homologs in *S. pombe*, *Plasmodium falciparum* and *H. sapiens* carry only one HIND element. And even more surprisingly, another splicing factor, Prp38p, carries C-terminal HIND element in *Plasmodium falciparum* and in *Arabidopsis thaliana* species. Functional consequences of this distribution of HIND elements in different species are not known (Mishra et al., 2011).

Binding of Hub1p to HIND is stronger than usual interaction of ubiquitin or UBLs with their receptors. Interactions of ubiquitin-UIMs (ubiquitin interaction motifs) are mostly hydrophobic in character and thus often weak, corresponding to $K_d \geq 100\ \mu M$ (Mishra et al., 2011). Binding of SUMO to SIM (Sumo interaction motif, which is a peptide as short as nine residues) is stronger, with K_d ranging between 5 μM and 10 μM (*Hochstrasser, 2009). Among these interactions, Hub1p binding to HIND-I is strongest, corresponding to $K_d \geq 1.69 \pm 0.27\ \mu M$ (Mishra et al., 2011). This tight binding is mostly caused by the presence of the salt bridge between aspartate-22 (D22) of Hub1p and arginine-16 (R16) of HIND-I (Figure 4b). In line with their importance, these residues are highly conserved among species, see Figure 2 (Karaduman et al., 2017; Mishra et al., 2011).

Structure of Hub1-HIND complex was compared with complexes of ubiquitin and SUMO with their binding domains. HIND element binds to the ubiquitin fold on the opposite side than UIM (Figure 4a). Despite the fact, that interaction surface for SIM is on the same side of the ubiquitin fold as the binding site for HIND, their two interaction interfaces are not identical (Mishra et al., 2011).

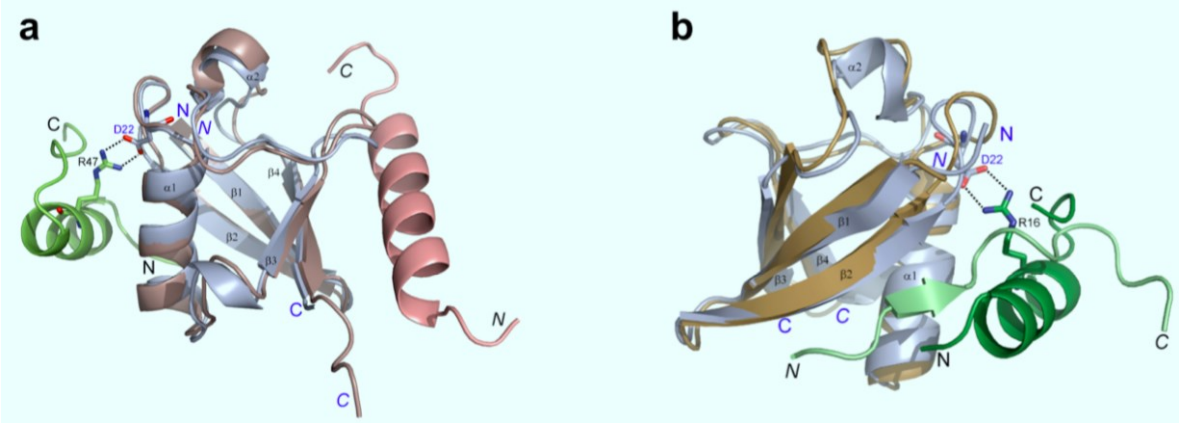

Figure 4: Comparison of binding patterns of complexes Hub1-HIND, ubiquitin-UIM, SUMO-SIM. (a) Superimposition of Hub1-HIND-II complex (grey and green respectively) and ubiquitin-UIM (brown and pink respectively). UIM binds to the ubiquitin β-grasp fold on the opposite side than HIND. **(b)** Superimposition of Hub1-HIND-I complex (grey and green respectively) and SUMO-SIM complex (dark yellow and light green respectively). Binding sites for SIM and HIND are right next to each other. HIND – Hub1 interaction domain, UIM – ubiquitin interaction motif, SIM – SUMO interaction motif. Figure taken from (Mishra et al., 2011).

4.2 Interaction with Prp5

Studies of spliceosome assembly *in vitro* showed that Hub1p associates stably with early spliceosome complexes E and pre-spliceosomes A assembled on substrates possessing nonconsensus 5'SS sequences (Karaduman et al., 2017). Moreover, in the absence of Hub1 protein, spliceosomes assembled on these suboptimal substrates stalled at the stage of H-complex (Karaduman et al., 2017) [note: H-complex is a nonessential spliceosomal stage, consisting of pre-mRNA and variable protein composition, it is the earliest stage of spliceosomal cycle going before early spliceosome E complex stage (*Coelho and Smith, 2014)]. Already known binding partner of Hub1p, Snu66p, was demonstrated to join the spliceosome later with tri-snRNP and thus it was not present already in E and A complexes. It means that another protein or possibly more proteins interact with Hub1p during early spliceosome assembly (Karaduman et al., 2017).

This interactor was identified as Prp5 by means of Sulfo-MBS chemical crosslinking [m-maleimidobenzoyl-N-hydroxysulfosuccinimide ester, with spacer arm size 7.3 Å (Thermo Fisher Scientific, accessed 2018)] and mass-spectrometry identification of isolated early spliceosomes. By mutagenesis of Hub1 protein, the contact surface with Prp5p was identified (Figure 5, residues identified as important for this interaction are depicted by red arrowheads also in Figure 2). This interface differs from the interaction surface for Snu66p and thus Prp5-contacting interface was termed Surface II and element interacting with HIND of Snu66p was termed Surface I. Histidine-63 of Surface II and tryptophan-257 of Prp5p were identified as critical residues for this interaction (Karaduman et al., 2017).

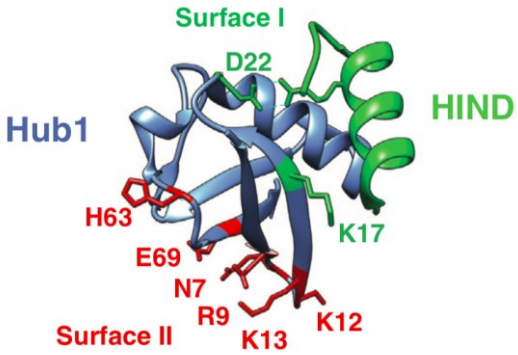

Figure 5: Positions of Surfaces I and II of Hub1 protein. Interaction interfaces of Prp5p (red) and Snu66p (green) are highlighted on the structure of Hub1p (blue) in complex with HIND element (green helix). Figure taken from (Karaduman et al., 2017).

5 FUNCTION OF HUB1 PROTEIN IN *S. CEREVISIAE*

5.1 Absence of Hub1p influences splicing of certain 5' splice sites

5.1.1 Identification of Hub1–dependent 5' splice sites by use of reporter systems

It was found out that splicing of certain types of introns in *S. cerevisiae* negatively reacts to the absence of Hub1p. This discovery was made by means of the 5'splice site mutagenesis of intron containing reporter fusion *RP51*-LacZ* (Mishra et al., 2011). Production of mature *LacZ* mRNA and thus β-galactosidase activity is dependent on proper splicing. Mutations of the 5'SS change the splicing efficiency and consequently the amount of β-galactosidase produced. By comparison of β-galactosidase activity of mutated sample and control carrying wild type (wt) reporter, relative quantification of splicing efficiency is acquired (Jacquier et al., 1985). Five nonconsensus 5'SS sequences on this reporter were found to be spliced inefficiently in the absence of Hub1p (Table 1). From now on, I will call these five sequences as Hub1-dependent, and all other nonconsensus 5'SS sequences as Hub1-independent. Additional five nonconsensus 5'SS were examined in *RP51*-LacZ* reporter system and their splicing did not exhibit any defects in the absence of Hub1p, see Table 1 (Mishra et al., 2011).

Hub1-dependent 5'SS	**Some Hub1-independent 5'SS**
GUAUAU	GCAUGU
GUCUGU	GUAUGC
GUCAGU	GUAAGU
GCAAGU	GUACGU
GUGAGU	GUAUGA
Consensus 5'SS sequence in *S. cerevisiae* is GUAUGU	

Table 1: Sequences of Hub1-dependent 5'SS and some Hub1-independent 5'SS. The first column shows sequences of 5'SS which demanded presence of Hub1 protein for splicing in reporter fusion *RP51*-LacZ* assay. In the second column 5'SS sequences of introns which were not affected by the absence of Hub1 protein in the same reporter system are given. Underlined nucleotides differ from the consensus sequence. Data obtained from (Mishra et al., 2011).

Real requirement of Hub1 protein for efficient splicing of GUAUAU and GUCUGU Hub1-dependent 5' splice sites was verified in yet another splicing-sensitive *ACT1-CUP1* reporter assay (Mishra et al., 2011). In brief, *CUP1* is nonessential gene ensuring survival of the cell when copper is present in the environment. Splicing efficiency of *ACT1-CUP1* fusion gene correlates with the degree of tolerance to higher copper concentrations (Lesser and Guthrie, 1993). Apart from verification, the use of *ACT1-CUP1* assay conferred some new findings. Some alterations of BP (UCCUAAC, UACUACC instead of UACUAAC) and 3'SS (AG/G, UG/U instead of AG/U, slash represents intron/exon boundary) were tested and splicing via these splicing signals was not affected in *hub1*Δ strains (Mishra et al., 2011).

5.1.2 Splicing-sensitive microarray

Splicing-sensitive microarray was used to obtain wide set of data about influence of Hub1p depletion on naturally occurring introns in *S. cerevisiae* (Mishra et al., 2011). In *S. cerevisiae* genome, only five introns are found whose 5'SS sequences are Hub1-dependent according to the above definition. Namely introns in genes *SPO22, RPL20A, SRC1* (upstream and downstream 5'SS) and *RPL30*, see Table 2 (Ares Lab Yeast Intron Database 4.3, accessed 2018). However, out of these introns, only *RPL30* and *SRC1* introns were present in the splicing-sensitive microarray testing set (Mishra et al., 2011). Microarray results showed only one defectively spliced intron in *hub1*Δ strain. Surprisingly, it was neither *RPL30* nor *SRC1*. It was *RPL34B,* which maybe even more surprisingly possesses the intron without Hub1-dependent 5'SS. (Mishra et al., 2011). Actually, all three splicing signals of *RPL34B* intron are canonical (Ares Lab Yeast Intron Database 4.3, accessed 2018).

Gene name	***SRC1* (upstream)**	***SRC1* (downstream)**	***RPL30***	***SPO22***	***RPL20A***
5'SS sequence	GCAAGU	GUGAGU	GUCAGU	GUAUAU	GUGAGU

Table 2: 5'SS sequences of Hub1-dependent introns. Five introns of *S. cerevisiae* contain Hub1-dependent 5'SS. *SRC1* has two overlapping Hub1-dependent 5'SS. Underlined nucleotides differ from the consensus sequence. Sequences of *SRC1* obtained from (Mishra et al., 2011) and sequence of *RPL30, SPO22* and *RPL20A* from (Ares Lab Yeast Intron Database 4.3, accessed 2018).

From these findings it seems that some introns naturally present in *S. cerevisiae* which include Hub1-dependent 5'SS do not react to the absence of Hub1 protein. Specifically, introns of *SRC1* and *RPL30* genes. Why splicing efficiency of these transcripts did not change in *hub1*Δ strain? Firstly, the reason why no alterations in splicing of *SRC1* were detected, lies in the structure of its 5'splice sites and the used method (Mishra et al., 2011).

In brief, splicing-sensitive microarray is based on the hybridization of total polyadenylated RNA with three types of oligonucleotide probes for every gene. The first is spliced-junction probe, which specifically reacts with spliced transcripts, the second probe is specific to introns and the third is a probe directed against the second exon of the gene and it is thus specific against both spliced and unspliced transcripts. By comparison of signals from these three probes in wt and mutated cells, the quantitative description of changes in splicing efficiency is derived (Clark et al., 2002). *SRC1* carries two overlapping 5'SS and as a result, spliced-junction probe reacts with both emerging transcripts. The assay described above is thus not able to distinguish between these two forms (Mishra et al., 2011). Despite the fact that total splicing efficiency is not affected by the absence of Hub1 protein, the ratio of products is affected. These two Hub1-dependent 5'SS therefore react very specifically to the absence of Hub1 protein and this regulation of *SRC1* alternative splicing is discussed in detail in the chapter 5.1.3.

Nevertheless, there is no obvious technical reason for no detected changes in level of *RPL30* on splicing-sensitive microarray, except for the possibility of the change being under the detection limit of this method. To confirm the microarray results, quantitative polymerase chain reaction with reverse transcription (RT-qPCR) was applied on 6 selected introns and again only splicing defect of *RPL34B* was confirmed. There was a higher proportion of unspliced transcripts in *RPL30* sample too, however, this result was not significant. It means that splicing efficiency of *RPL30* also reacts to the absence of Hub1p, but only slightly (Mishra et al., 2011).
In summary, both *RPL30* and *SRC1*, in spite of being negative on splicing-sensitive microarray, are somehow affected by the absence of Hub 1p. It would be highly interesting to see whether splicing efficiency of *SPO22* and *RPL20A* introns (which possess Hub1-dependent 5'SS as well but were not present in the microarray set) does react to the Hub1-depletion and to what extent.

The only intron significantly sensitive to the absence of Hub1p in microarray assay was *RPL34B* without Hub1-dependent 5'SS (Mishra et al., 2011). Based on this, I might speculate that possession of Hub1-dependent 5'SS sequence is not the single condition for creation of dependence in splicing efficiency on Hub1p and that other factors are probably important.
Splice site recognition is not dependent only on sequence information of splice sites, but it is also affected by the quality of surrounding sequences. Neighboring sequences may create RNA secondary structures, which would interfere with splice site recognition. The most obvious example would be a sequestration of a splice site sequence and its resulting unavailability to snRNAs and spliceosomal proteins (*Coelho and Smith, 2014). Naturally occurring introns, in contrast to introns of the reporter systems, are set in a different structural context and this might be an explanation for contradiction between sensitivity of *RPL34B* in microarray and Hub1-independence of its 5'SS determined by reporter system. Hub1-dependence might arise from other pre-mRNA features apart from sequence information, secondary RNA structure being just one example, and this might be the case of *RPL34B*.

5.1.3 Hub1 protein regulates alternative splicing of *SRC1*

SRC1 is one of the handful of genes in *S cerevisiae*, that are subject to alternative splicing which gives rise to two distinct protein products with different functions (Schreiber et al., 2015, Grund et al., 2008) and alternative splicing of this gene is affected by the absence of Hub1p (Mishra et al., 2011). Structure of 5'splice sites of a single intron of *SRC1* is depicted in Figure 6.

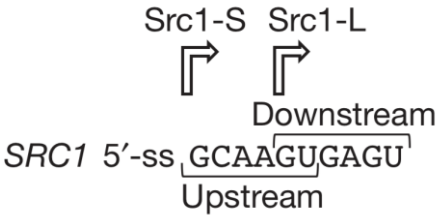

Figure 6: Overlapping 5' splice sites of *SRC1*. Intron of *SRC1* possesses two overlapping 5' splice sites which are Hub1-dependent by the sequence. Splicing via these sites gives rise to two distinct transcripts, major Src1-S and minor Src1-L. Modified from (Mishra et al., 2011).

In wt cells the ration of transcripts Src1-S/Src1-L equals to 60/40 but in *hub1*Δ three-fold decrease in Src1-S level is observed (Kawashima et al., 2014). Each 5'SS of *SRC1* is Hub1-dependent by the sequence, and factual decrease of splicing efficiency in *hub1*Δ cells was confirmed when each of these sites was separately tested in *RP51*-LacZ* reporter system. However, in natural overlapping arrangement, only the efficiency of splicing via upstream 5'SS is decreased when Hub1p is not present. In *hub1*Δ cells Src1-S transcript is almost absent whereas level of Src1-L is normal or even slightly increased. These findings were obtained by cDNA sequencing across exon/exon boundary and visualization of tagged spliced variants on electrophoretic gel (Mishra et al., 2011).

Interesting consequences result from mutations of the 5' splice sites. Certain mutations in the upstream 5'SS provide Hub1-independence of splicing vi this site but simultaneously they cause complete repression of the splicing via downstream 5'SS. Similarly, mutations of the downstream 5'SS affect splicing via upstream 5'SS. Different combinations of mutations change the ratio of Src1-S/Src1-L in various ways. Considering all this, overlapping arrangement of splice sites seems to be finely tuned mechanism for keeping appropriate proportion of two spliced products and Hub1p at least participate on keeping this balance (Mishra et al., 2011).

5.2 Overexpression of Hub1 protein impacts splicing

I have reported various effects of Hub1 depletion and now I will turn my attention to consequences of Hub1p overexpression, which was assessed by RNA sequencing. Excessive amount of this protein causes missplicing and it was confirmed that HUB-OE (Hub1p overexpression) is toxic for cells with unfunctional nonsense mediated decay and RNA surveillance mechanism. In order to preserve misspliced products for RNA sequencing analysis, effects of overexpression were examined in the *xrn1Δ rat1-1* mutants (Karaduman et al., 2017).

RNA sequencing brought interesting results. Splicing of 43 introns is enhanced in HUB1-OE strain, but simultaneously 16 events of cryptic splicing are induced. A common attribute of the 43 enhanced introns is deviation from consensus on certain positions of 5'SS and BP sequence (Figure 7a). Specifically, the 3rd, the 4th and the 6th positions of 5'SS and the 1st position of BP sequence manifested lowered representation of consensus nucleotides. Other parameters of introns - intron length, level of expression, alterations of 3'SS from consensus, GC-content, pyrimidine content, distance between 5'SS and BP, distance between BP and 3'SS – were not significantly altered. Distribution of sequences of 16 cryptic splicing events is displayed in the pie chart (Figure 7b). Majority of these cryptic splicing events is mediated through various combinations of nonconsensus splicing signals, but surprisingly, about one third of events is implemented via consensus splicing signals (Karaduman et al., 2017).

Change in splicing of *RPL34B* under overexpression of Hub1p was not detected (Karaduman et al., 2017). Neither was detected any influence on the splicing level of *SRC1* or *RPL30* genes, which possess Hub1-dependent 5'SS (Karaduman et al., 2017).

To summarize the results, it seems that splicing of *RPL34B* is sensitive to low levels of Hub1p (Mishra et al., 2011) but does not react to high levels (Karaduman et al., 2017). Despite the fact, that the ratio of Src1-S/Src1-L is altered when Hub1p is absent (Mishra et al., 2011), it does not change under Hub1p overexpression (Karaduman et al., 2017). And finally, out of 43 HUB1-OE activated introns (Karaduman et al., 2017), splicing of 27 is not affected in *hub1*Δ and residual 16 were not present on the microarray in *hub1*Δ (Mishra et al., 2011). Apart from *RPL34B* and *SRC1* alternative splicing regulation it seems that for most of the affected introns, Hub1p is not strictly required, but might function as an activator of their splicing if abundant (Karaduman et al., 2017). However, it would be interesting to see if Hub1p depletion has any effect on residual 16 HUB1-OE-activated introns.

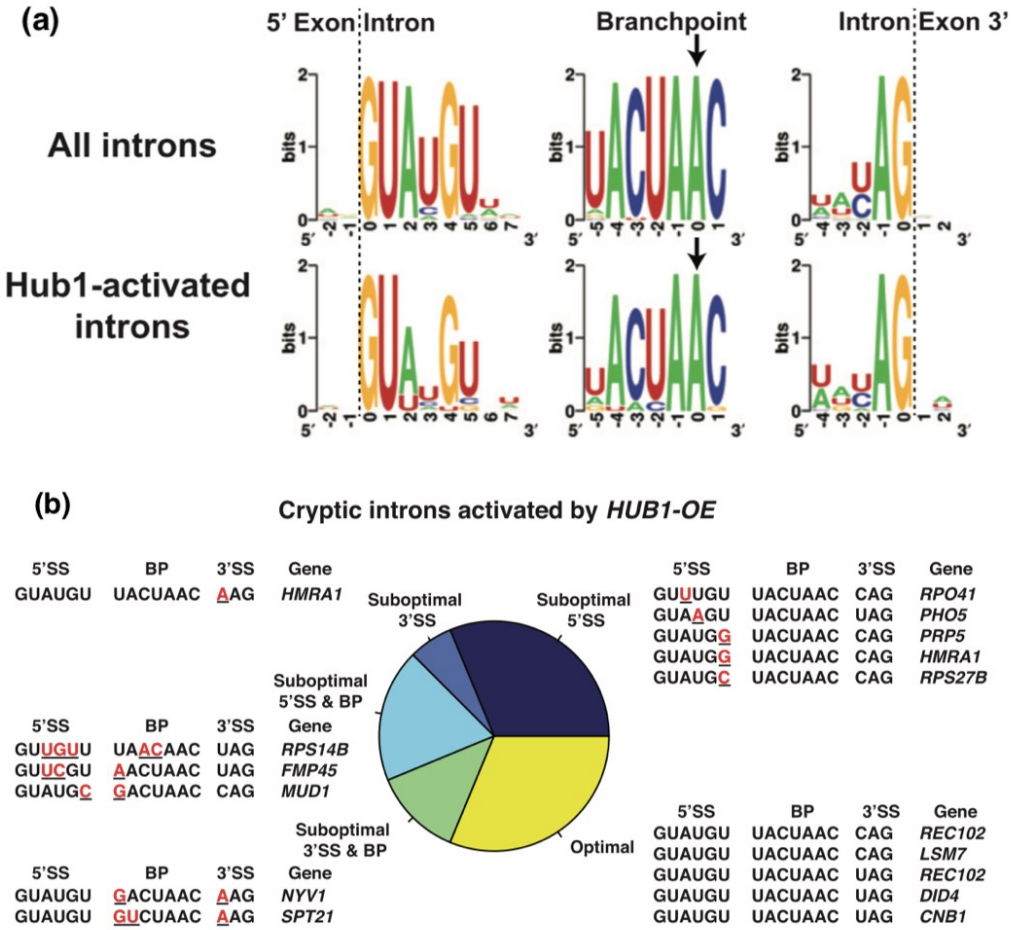

Figure 7: (a) 43 Hub1-activated introns display variation from consensus 5'SS a BP sequence. (b) Distribution of 16 cryptic splicing signals. (a) Splicing signals of 43 Hub1-activated introns vary from the consensus mainly on the 3rd, the 4th and the 6th position of 5'SS and the 1st position of BP. (b) About a third of the cryptic splicing events is connected with altered 5'SS and approximately one quarter is linked to combinations of nonconsensus 5'SS and BP. Minority is assigned to alterations of 3'SS or combination of altered 3'SS and BP. Figure modified from (Karaduman et al., 2017)

5.2.1 Specific type of mistake induced by Hub1 protein overexpression

Specific substrate carrying optimal 5'SS was found to be error-inducing in HUB-OE but not in *hub1*Δ or wt cell extracts. This specific substrate consists of optimal 5'SS and 4 upstream guanosines: GˇGˇGˇGˇGUAUGU. By primer extension analysis, missplicing was identified to happen via guanosines upstream of canonical 5'SS (red arrowheads designate the sites of missplicing) (Karaduman et al., 2017).

5.3 Hub1 protein seems to have two faces

On one hand, overexpression of Hub1 protein might have positive effect on the cell since it enhances splicing of 43 introns, but on the other hand this state also causes 16 events of cryptic splicing. (Karaduman et al., 2017). The important question is, what significance this error-prone mechanism might have for the cell.

Karaduman et al. (2017) claim that "error-prone mechanisms in the face of obstacles are common in nature" and give an example of translesion polymerases. Constitutive splicing hand in hand with cryptic splicing enhanced by Hub1 protein might be a mechanism of this type. However, there is another possible point of view. Sixteen cryptic splicing events not necessarily have to be accidental. One striking example is the cryptic splicing of *prp5* transcript, which regulates level of Prp5p and serves as a safeguarding mechanism against missplicing induced by Hub1 protein overactivity (Karaduman et al., 2017).

5.3.1 Negative feedback loop

Apart from the strong association of Hub1p with Prp5p in early spliceosomal complexes assembled on suboptimal substrates, Hub1p operates as an ATPase-activity stimulating cofactor of Prp5 helicase (Karaduman et al., 2017). As can be seen on the structure of Prp5p in complex with ADP (Zhang et al., 2013), the binding site of ATP is near tryptophan-257, which is the residue critical for interaction with Surface II of Hub1 (Figure 8). *In vitro* experiments revealed that RNA-dependent hydrolytic activity of recombinant Prp5p is enhanced after addition of Hub1p. Conversely addition of Hub1-H63L protein, the mutant of Surface II which is defective in Prp5p binding, did not enhance ATPase activity, which demonstrates the relevance of Surface II of Hub1p for this enhancement (Karaduman et al., 2017).

One of the 16 Hub1-induced cryptic splicing events targets on *prp5* transcript. Cryptic intron is removed through the suboptimal 5'SS GUAUGG which contains the start codon and thus this triplet is consequently removed with intron (start codon is designated in red, nucleotides differing from the consensus sequence are underlined). In this way, Hub1-induced cryptic splicing prevents the emergence of essential Prp5 helicase (Karaduman et al., 2017).

In HUB1-OE cells the level of cryptically-spliced transcript is high, whereas in wt cells the level is low, and it is hardly detectable in *hub1*Δ cells. Observed levels of Prp5 protein are in agreement with transcript levels. Specifically, the lowest level in HUB1-OE (0.6 of the wt level) and the highest level in *hub1*Δ (1.6 of the wt level) (Karaduman et al., 2017).

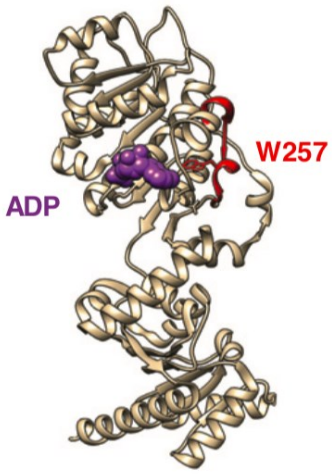

Figure 8: Structure of Prp5 helicase in complex with ADP. Position of tryptophan-257, the important residue for Hub1 binding, is highlighted. Original structure of Prp5-ADP complex was published by (Zhang et al., 2013). Figure was taken from (Karaduman et al., 2017).

The important part of the experimental background was monitoring of the toxicity level conferred by overexpression of Hub1p in cells with defective nonsense mediated decay and RNA surveillance mechanism (xrn1Δ rat1-1 mutants). When Hub1 protein is overexpressed in these defective cells and simultaneously a nonspliceable variant $Prp5^{NS}$ (constructed by change of the cryptic 5'SS from the wt GUAUGG to AUAUGG) is expressed, a measured toxicity rises. However, when $Prp5^{NS}$ is expressed in these cells alone, the toxicity is comparable with the wt cells (xrn1Δ rat1-1 background). It means that uncontrollability of Prp5 splicing through Hub1-induced cryptic splicing increases the number of aberrantly spliced transcripts, which are the primary cause of toxicity. Under normal conditions, protein level of wt Prp5 would react to an elevated level of Hub1p by its own decrease and in this way, harmful impacts of Hub1p overactivity would be reduced. This mechanism functions as a negative feedback loop safeguarding the cell from the negative impact of Hub1p hyperactivity (Karaduman et al., 2017).

Negative-feedback loop regulating Prp5p level according to the actual amount of Hub1p is the example of a cryptic splicing utilization for regulation purposes. I dare to speculate, that possibly other 15 cryptic splicing events, or at least some of them, could be functional as well rather than erroneous. However, observation that *hub1* transcripts are inducible by cadmium and oxidative stress testifies in favour of a concept of Hub1p as an error-prone activator used to overcome obstacles. Nevertheless, non *hub1*Δ-specific defects were detected under exposition to these stress conditions or heat shock (Mishra et al., 2011). Also gene ontology analysis of Hub1-activated introns did not show a significant enrichment of genes from any particular cellular pathway (Karaduman et al., 2017). However, some interesting observations in this respect were made when a water stress resistance in perennial ryegrass was studied (Patel et al., 2015).

5.3.2 Water-stress response mediated by Hub1 homolog in *Lolium perenne*

Patel et al. (2015) assessed the involvement of LpHUB1 gene in perennial ryegrass (*Lolium perenne L.)* while exposed to a water stress. On the basis of previously observed expression patterns, which registered elevated level of LpHUB1 products during summer heat and under water stress, they speculated that the overexpression of LpHUB1 might improve resistance to water stress and tested this hypothesis experimentally. A better growth and improved turf quality were indeed observed in plants comprising transgenic LpHUB1 after long-term exposure to the insufficient water supplies (Patel et al., 2015).

Knowing the impacts of Hub1p overexpression on splicing in *S. cerevisiae* and presuming the conservation of Hub1p function in eukaryotes, it is speculated that Hub1p might rescue stress-related splicing defects and thus increase water-stress tolerance. Hub1 protein thus might be in the end the error-prone activator of splicing, which in the face of stress rescues impaired splicing, but with the drawback of causing cryptic splicing errors (Karaduman et al., 2017).

5.4 Hub1-dependent splicing in other organisms

Whether the mechanism of function of Hub1p is conserved among eukaryotes is not currently clear. However, research was made in *S. pombe* and human cells and results suggest that general involvement of Hub1 protein in pre-mRNA splicing modulation might be widespread (Ammon et al., 2014; Mishra et al., 2011).

5.4.1 Hub1-dependent splicing in *S. pombe*

In contrast to *S. cerevisiae*, *HUB1* gene in *S. pombe* is essential (Wilkinson et al., 2004; Yashiroda and Tanaka, 2004). Efficient selective splicing of at least two genes of *S. pombe* was showed to depend on the presence of Hub1p (only two genes were tested altogether) (Mishra et al., 2011). *cdc2* and *zas1* genes have 5 and 8 introns respectively (PomBase, accessed 2018) and splicing of some of these introns is impaired in *hub1-1* temperature sensitive mutant (Mishra et al., 2011).

5.4.2 UBL5-dependent splicing in human cells

UBL5 protein in human cells is essential, in contrast to *S. cerevisiae* Hub1p (Ammon et al., 2014; Oka et al., 2014). UBL5 protein is known to colocalize with Cajal bodies (Švéda et al., 2013) which are subnuclear compartments where snRNA modifications and snRNP assembly takes place (*Herzel et al., 2017) which suggests functional link of *UBL5* to splicing. Later UBL5p depletion was observed to cause a severe cell cycle progression delay, chromosome segregation defects and eventually apoptosis (Ammon et al., 2014; Oka et al., 2014).

Intriguingly, a premature loss of sister chromatid cohesion in *UBL5* knockdown cells was observed and Oka et al. (2014) managed to identify the underlaying, splicing-related mechanism. UBL5p depletion enhances intron retention in the transcript of a key cohesion factor Sororin, which results in the lowered Sororin protein level and consequently in chromatid cohesion defects (Oka et al., 2014).

Apart from this specific splicing-related activity, *UBL5* is required for appropriate splicing of many human introns. More than 3 000 of altered splicing events were identified in UBL5p depleted cells (Ammon et al., 2014). However, some introns are not affected by the absence of UBL5p at all. Even in the single affected pre-mRNA only a portion of introns is sensitive to the UBL5p depletion. Absence of UBL5 protein results in various types of splicing defects: exon skipping, lower steady-state mRNA level and intron retention. Otherwise splice sites of influenced introns show no sequence similarity (Ammon et al., 2014).

It seems that the role of UBL5 protein in pre-mRNA splicing regulation is evolutionary conserved. (Ammon et al., 2014; Oka et al., 2014).

5.5 Model of Hub1 function during splicing reaction

Initial concept of Hub1-dependent 5'splice sites was the important step, but following experiments showed that a wide spectrum of introns with diverse splicing signals reacts to an elevated level of Hub1 protein. Eventually, on the basis of observed interaction with Prp5 helicase, a general model of Hub1p function during the splicing reaction was proposed (Karaduman et al., 2017).

It is known that splice sites recognition is influenced by DExD/H-box helicases and that specifically Prp5 helicase directs early stages of spliceosome assembly (*Shi, 2017). According to the model, Prp5 ATPase activity is induced by Hub1p which relaxes its stringency and allows selection of suboptimal splice sites. However, when the level of Hub1p is low, it associates mainly with Snu66p and for suboptimal splice sites, the level of Prp5-ATPase-activity stimulation by the rest of Hub1p is not sufficient. Conversely, when the level of Hub1p is high, Prp5 helicase is hyperactivated which results not only in the usage of suboptimal splice sites but also in overall enhancement of splicing rate and progression of spliceosome assembly. Thus Hub1-level-dependent activation of Prp5 affects both splice-site types, optimal and suboptimal. In this model, Hub1p fulfils the role of error-prone splicing stimulator with impact on splicing fidelity by utilization of suboptimal splice sites (Karaduman et al., 2017).

6 CONCLUSION

This literary summary gives a detailed description of Hub1 protein and in the end, it provides an overview of Hub1 as a representative of an elegant splicing regulation mechanism. Noncovalent interaction of a small protein with a core spliceosomal component modulates a splice site utilization and thus possibly promotes a stress-related response (Karaduman et al., 2017).

But what effect has the interaction of Snu66 with Hub1 on the splicing process? What are the pathways that regulate the emergence of Hub1 protein itself? Is this regulation mechanism used widely?

Many questions about Hub1 protein remain to be answered.

7 REFERENCES

Ammon, T., Mishra, S.K., Kowalska, K., Popowicz, G.M., Holak, T.A., Jentsch, S., 2014. The conserved ubiquitin-like protein Hub1 plays a critical role in splicing in human cells. J. Mol. Cell Biol. 6, 312–323. https://doi.org/10.1093/jmcb/mju026

Bai, R., Yan, C., Wan, R., Lei, J., Shi, Y., 2017. Structure of the Post-catalytic Spliceosome from Saccharomyces cerevisiae. Cell 171, 1589-1598.e8. https://doi.org/10.1016/j.cell.2017.10.038

Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R., Becker, J., 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. J. Mol. Biol. 280, 275–286. https://doi.org/10.1006/jmbi.1998.1839

Bertram, K., Agafonov, D.E., Dybkov, O., Haselbach, D., Leelaram, M.N., Will, C.L., Urlaub, H., Kastner, B., Lührmann, R., Stark, H., 2017. Cryo-EM Structure of a Pre-catalytic Human Spliceosome Primed for Activation. Cell 170, 701-713.e11. https://doi.org/10.1016/j.cell.2017.07.011

Clark, T.A., Sugnet, C.W., Ares, M., 2002. Genomewide analysis of mRNA processing in yeast using splicing-specific microarrays. Science 296, 907–910. https://doi.org/10.1126/science.1069415

Coelho, M.B., Smith, C.W.J., 2014. Regulation of alternative pre-mRNA splicing. Methods Mol. Biol. Clifton NJ 1126, 55–82. https://doi.org/10.1007/978-1-62703-980-2_5

Collier, G.R., McMillan, J.S., Windmill, K., Walder, K., Tenne-Brown, J., de Silva, A., Trevaskis, J., Jones, S., Morton, G.J., Lee, S., Augert, G., Civitarese, A., Zimmet, P.Z., 2000. Beacon: a novel gene involved in the regulation of energy balance. Diabetes 49, 1766–1771.

Deckert, J., Hartmuth, K., Boehringer, D., Behzadnia, N., Will, C.L., Kastner, B., Stark, H., Urlaub, H., Lührmann, R., 2006. Protein Composition and Electron Microscopy Structure of Affinity-Purified Human Spliceosomal B Complexes Isolated under Physiological Conditions. Mol. Cell. Biol. 26, 5528–5543. https://doi.org/10.1128/MCB.00582-06

Dittmar, G.A.G., Wilkinson, C.R.M., Jedrzejewski, P.T., Finley, D., 2002. Role of a ubiquitin-like modification in polarized morphogenesis. Science 295, 2442–2446. https://doi.org/10.1126/science.1069989

Friedman, J.S., Koop, B.F., Raymond, V., Walter, M.A., 2001. Isolation of a ubiquitin-like (UBL5) gene from a screen identifying highly expressed and conserved iris genes. Genomics 71, 252–255. https://doi.org/10.1006/geno.2000.6439

Galej, W.P., Wilkinson, M.E., Fica, S.M., Oubridge, C., Newman, A.J., Nagai, K., 2016. Cryo-EM structure of the spliceosome immediately after branching. Nature 537, 197–201. https://doi.org/10.1038/nature19316

Grund, S.E., Fischer, T., Cabal, G.G., Antúnez, O., Pérez-Ortín, J.E., Hurt, E., 2008. The inner nuclear membrane protein Src1 associates with subtelomeric genes and alters their regulated gene expression. J. Cell Biol. 182, 897–910. https://doi.org/10.1083/jcb.200803098

Jacquier, A., Rodriguez, J.R., Rosbash, M., 1985. A quantitative analysis of the effects of 5' junction and TACTAAC box mutants and mutant combinations on yeast mRNA splicing. Cell 43, 423–430.

Karaduman, R., Chanarat, S., Pfander, B., Jentsch, S., 2017. Error-Prone Splicing Controlled by the Ubiquitin Relative Hub1. Mol. Cell 67, 423-432.e4. https://doi.org/10.1016/j.molcel.2017.06.021

Kawashima, T., Douglass, S., Gabunilas, J., Pellegrini, M., Chanfreau, G.F., 2014. Widespread Use of Non-productive Alternative Splice Sites in Saccharomyces cerevisiae. PLOS Genet. 10, e1004249. https://doi.org/10.1371/journal.pgen.1004249

Lesser, C.F., Guthrie, C., 1993. Mutational analysis of pre-mRNA splicing in Saccharomyces cerevisiae using a sensitive new reporter gene, CUP1. Genetics 133, 851–863.

Li, Z., Lee, I., Moradi, E., Hung, N.-J., Johnson, A.W., Marcotte, E.M., 2009. Rational extension of the ribosome biogenesis pathway using network-guided genetics. PLoS Biol. 7, e1000213. https://doi.org/10.1371/journal.pbio.1000213

Liang, W.-W., Cheng, S.-C., 2015. A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence. Genes Dev. 29, 81–93. https://doi.org/10.1101/gad.253708.114

Lüders, J., Pyrowolakis, G., Jentsch, S., 2003. The ubiquitin-like protein HUB1 forms SDS-resistant complexes with cellular proteins in the absence of ATP. EMBO Rep. 4, 1169–1174. https://doi.org/10.1038/sj.embor.7400025

McNally, T., Huang, Q., Janis, R.S., Liu, Z., Olejniczak, E.T., Reilly, R.M., 2003. Structural analysis of UBL5, a novel ubiquitin-like modifier. Protein Sci. Publ. Protein Soc. 12, 1562–1566. https://doi.org/10.1110/ps.0382803

Mishra, S.K., Ammon, T., Popowicz, G.M., Krajewski, M., Nagel, R.J., Ares, M., Holak, T.A., Jentsch, S., 2011. Role of the ubiquitin-like protein Hub1 in splice-site usage and alternative splicing. Nature 474, 173–178. https://doi.org/10.1038/nature10143

Nguyen, T.H.D., Galej, W.P., Bai, X.-C., Oubridge, C., Newman, A.J., Scheres, S.H.W., Nagai, K., 2016. Cryo-EM structure of the yeast U4/U6.U5 tri-snRNP at 3.7 Å resolution. Nature 530, 298–302. https://doi.org/10.1038/nature16940

Oka, Y., Varmark, H., Vitting-Seerup, K., Beli, P., Waage, J., Hakobyan, A., Mistrik, M., Choudhary, C., Rohde, M., Bekker-Jensen, S., Mailand, N., 2014. UBL5 is essential for pre-mRNA splicing

and sister chromatid cohesion in human cells. EMBO Rep. 15, 956–964. https://doi.org/10.15252/embr.201438679

Patel, M., Milla-Lewis, S., Zhang, W., Templeton, K., Reynolds, W.C., Richardson, K., Biswas, M., Zuleta, M.C., Dewey, R.E., Qu, R., Sathish, P., 2015. Overexpression of ubiquitin-like LpHUB1 gene confers drought tolerance in perennial ryegrass. Plant Biotechnol. J. 13, 689–699. https://doi.org/10.1111/pbi.12291

Philippsen, P., Kleine, K., Pöhlmann, R., Düsterhöft, A., Hamberg, K., Hegemann, J.H., Obermaier, B., Urrestarazu, L.A., Aert, R., Albermann, K., Altmann, R., André, B., Baladron, V., Ballesta, J.P., Bécam, A.M., Beinhauer, J., Boskovic, J., Buitrago, M.J., Bussereau, F., Coster, F., Crouzet, M., D'Angelo, M., Dal Pero, F., De Antoni, A., Del Rey, F., Doignon, F., Domdey, H., Dubois, E., Fiedler, T., Fleig, U., Floeth, M., Fritz, C., Gaillardin, C., Garcia-Cantalejo, J.M., Glansdorff, N.N., Goffeau, A., Gueldener, U., Herbert, C., Heumann, K., Heuss-Neitzel, D., Hilbert, H., Hinni, K., Iraqui Houssaini, I., Jacquet, M., Jimenez, A., Jonniaux, J.L., Karpfinger, L., Lanfranchi, G., Lepingle, A., Levesque, H., Lyck, R., Maftahi, M., Mallet, L., Maurer, K.C., Messenguy, F., Mewes, H.W., Mösti, D., Nasr, F., Nicaud, J.M., Niedenthal, R.K., Pandolfo, D., Piérard, A., Piravandi, E., Planta, R.J., Pohl, T.M., Purnelle, B., Rebischung, C., Remacha, M., Revuelta, J.L., Rinke, M., Saiz, J.E., Sartorello, F., Scherens, B., Sen-Gupta, M., Soler-Mira, A., Urbanus, J.H., Valle, G., Van Dyck, L., Verhasselt, P., Vierendeels, F., Vissers, S., Voet, M., Volckaert, G., Wach, A., Wambutt, R., Wedler, H., Zollner, A., Hani, J., 1997. The nucleotide sequence of Saccharomyces cerevisiae chromosome XIV and its evolutionary implications. Nature 387, 93–98.

Plaschka, C., Lin, P.-C., Nagai, K., 2017. Structure of a pre-catalytic spliceosome. Nature 546, 617–621. https://doi.org/10.1038/nature22799

Ramelot, T.A., Cort, J.R., Yee, A.A., Semesi, A., Edwards, A.M., Arrowsmith, C.H., Kennedy, M.A., 2003. Solution structure of the yeast ubiquitin-like modifier protein Hub1. J. Struct. Funct. Genomics 4, 25–30.

Sathasivam, K., Neueder, A., Gipson, T.A., Landles, C., Benjamin, A.C., Bondulich, M.K., Smith, D.L., Faull, R.L.M., Roos, R.A.C., Howland, D., Detloff, P.J., Housman, D.E., Bates, G.P., 2013. Aberrant splicing of HTT generates the pathogenic exon 1 protein in Huntington disease. Proc. Natl. Acad. Sci. U. S. A. 110, 2366–2370. https://doi.org/10.1073/pnas.1221891110

Schreiber, K., Csaba, G., Haslbeck, M., Zimmer, R., 2015. Alternative Splicing in Next Generation Sequencing Data of Saccharomyces cerevisiae. PLoS ONE 10. https://doi.org/10.1371/journal.pone.0140487

Stevens, S.W., Abelson, J., 1999. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. Proc. Natl. Acad. Sci. U. S. A. 96, 7226–7231.

Stevens, S.W., Barta, I., Ge, H.Y., Moore, R.E., Young, M.K., Lee, T.D., Abelson, J., 2001. Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from Saccharomyces cerevisiae. RNA N. Y. N 7, 1543–1553.

Švéda, M., Častorálová, M., Lipov, J., Ruml, T., Knejzlík, Z., 2013. Human UBL5 protein interacts with coilin and meets the Cajal bodies. Biochem. Biophys. Res. Commun. 436, 240–245. https://doi.org/10.1016/j.bbrc.2013.05.083

Vijay-Kumar, S., Bugg, C.E., Cook, W.J., 1987. Structure of ubiquitin refined at 1.8 A resolution. J. Mol. Biol. 194, 531–544.

Wan, R., Yan, C., Bai, R., Huang, G., Shi, Y., 2016. Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. Science 353, 895–904. https://doi.org/10.1126/science.aag2235

Wan, R., Yan, C., Bai, R., Lei, J., Shi, Y., 2017. Structure of an Intron Lariat Spliceosome from Saccharomyces cerevisiae. Cell 171, 120-132.e12. https://doi.org/10.1016/j.cell.2017.08.029

Whitby, F.G., Xia, G., Pickart, C.M., Hill, C.P., 1998. Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. J. Biol. Chem. 273, 34983–34991.

Wilkinson, C.R.M., Dittmar, G.A.G., Ohi, M.D., Uetz, P., Jones, N., Finley, D., 2004. Ubiquitin-like protein Hub1 is required for pre-mRNA splicing and localization of an essential splicing factor in fission yeast. Curr. Biol. CB 14, 2283–2288. https://doi.org/10.1016/j.cub.2004.11.058

Yan, C., Wan, R., Bai, R., Huang, G., Shi, Y., 2017. Structure of a yeast step II catalytically activated spliceosome. Science 355, 149–155. https://doi.org/10.1126/science.aak9979

Yan, C., Wan, R., Bai, R., Huang, G., Shi, Y., 2016. Structure of a yeast activated spliceosome at 3.5 Å resolution. Science 353, 904–911. https://doi.org/10.1126/science.aag0291

Yashiroda, H., Tanaka, K., 2004. Hub1 is an essential ubiquitin-like protein without functioning as a typical modifier in fission yeast. Genes Cells Devoted Mol. Cell. Mech. 9, 1189–1197. https://doi.org/10.1111/j.1365-2443.2004.00807.x

Zhang, X., Yan, C., Zhan, X., Li, L., Lei, J., Shi, Y., 2018. Structure of the human activated spliceosome in three conformational states. Cell Res. 28, 307–322. https://doi.org/10.1038/cr.2018.14

Zhang, Z.-M., Yang, F., Zhang, J., Tang, Q., Li, J., Gu, J., Zhou, J., Xu, Y.-Z., 2013. Crystal structure of Prp5p reveals interdomain interactions that impact spliceosome assembly. Cell Rep. 5, 1269–1278. https://doi.org/10.1016/j.celrep.2013.10.047

*Herzel, L., Ottoz, D.S.M., Alpert, T., Neugebauer, K.M., 2017. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. 18, 637–650. https://doi.org/10.1038/nrm.2017.63

*Hochstrasser, M., 2009. Origin and function of ubiquitin-like proteins. Nature 458, 422–429. https://doi.org/10.1038/nature07958

*Chiou, N.-T., Lynch, K.W., 2014. Mechanisms of spliceosomal assembly. Methods Mol. Biol. Clifton NJ 1126, 35–43. https://doi.org/10.1007/978-1-62703-980-2_3

*Newman, A.J., 2008. RNA Interactions in mRNA Splicing, in: John Wiley & Sons, Ltd (Ed.), Encyclopedia of Life Sciences. John Wiley & Sons, Ltd, Chichester, UK. https://doi.org/10.1002/9780470015902.a0000881.pub2

*Semlow, D.R., Staley, J.P., 2012. Staying on message: ensuring fidelity in pre-mRNA splicing. Trends Biochem. Sci. 37, 263–273. https://doi.org/10.1016/j.tibs.2012.04.001

*Shi, Y., 2017. The Spliceosome: A Protein-Directed Metalloribozyme. J. Mol. Biol., John Kendrew's 100th Anniversary Special Edition 429, 2640–2653. https://doi.org/10.1016/j.jmb.2017.07.010

*review

ONLINE RESOURCES AND DATABASES

Ares Lab Yeast Intron Database 4.3 (Nov 2011). http://intron.ucsc.edu/yeast4.3/ (accessed May 05, 2018).

PomBase. https://www.pombase.org/gene/SPBC11B10.09 (accessed May 02, 2018).

PomBase. https://www.pombase.org/gene/SPBC1198.04c (accessed May 02, 2018).

Saccharomyces genome database. https://www.yeastgenome.org/locus/S000007251/protein (accessed May 02, 2018).

Thermo Fisher Scientific. https://www.thermofisher.com/order/catalog/product/22312 (accessed May 05, 2018).