Charles University in Prague
Faculty of Science

PhD Thesis Summary

Regulation of pre-mRNA splicing in *S. cerevisiae*: where RNA cooperates with proteins

Ondřej Gahura
Prague 2011

Doctorate study programs in biomedicine
Charles University in Prague
and
Academy of Sciences of the Czech Republic

Program
Molecular and Cell Biology, Genetics and Virology

Board director
Prof. RNDr. Stanislav Zadražil, DrSc.

Training institution
Department of Cell Biology, Faculty of Science, Charles University in Prague

Author
Ondřej Gahura

Supervisor
Doc. RNDr. František Půta, CSc.

Contents

Abstract

Removal of introns from protein coding transcripts occurs in two splicing reactions catalyzed by a large nuclear complex, spliceosome. The spliceosome is an extremely intricate and dynamic machine, wherein contributions of small RNA molecules and multiple proteins are coordinated to meet the requirements of absolute precision and high flexibility. For an intimate understanding of pre-mRNA splicing, it is necessary to unravel roles of individual components and to dissect the partial mechanisms.

In the first part of this work, we describe the role of the Prp45 splicing factor in *Saccharomyces cerevisiae*. Mapping of genetic interactions of a conditionally lethal allele *prp45*(1-169) suggests a relationship of Prp45 to the NTC complex and to the second transesterification. Two-hybrid assay and purification of spliceosomal complexes reveal a contribution of the Prp45 C-terminus in the Prp22 helicase recruitment and/or regulation. Numerous experiments with reporter substrates document the need of Prp45 for the efficient splicing of a specific subset of introns. Our observations suggest that the function of Prp45 in splicing is conserved in evolution.

The second part is devoted to the role of intron secondary structure in 3´ splice site (3´ss) recognition. We show that the stem-loop structures formed downstream of the branch point (BP) are required for the splicing of *COF1* and *UBC13* introns, which have extremely long distances between BP and 3´ss. Identified structures aid to efficient 3´ss recognition by bringing remote 3´ss to the BP proximity and by sequestering AG dinucleotides, which behave as potential cryptic 3´ss. Our analyses strongly suggest that the structure based mechanism of 3´ss recognition is employed in most introns with distant 3´ss in *Saccharomyces cerevisiae* and possibly in other *Saccharomycotina* yeasts.

Introduction

The excision of introns from pre-mRNAs by spliceosome is one of the prerequisites for production of messenger RNA translatable into functional proteins. Despite the apparent chemical simplicity of the reaction, splicing represents one of the most intricate cellular processes. To complement the complexity of the machinery, numerous associated proteins link splicing with preceding and subsequent stages of gene expression (*1, 2*). The spliceosome assembles on pre-mRNA in a stepwise manner. During whole process multiple rearrangements within spliceosome occur, including repeated reorganizations of its catalytic center. Transitions between particular stages are associated with the release of specific proteins and with the recruitment of others (reviewed in (*3*)).

Roles of numerous splicing factors are rather ambiguous. In this work, we present the study on the role of the essential protein Prp45 in *Saccharomyces cerevisiae*. In budding yeast, it was implicated only in splicing (*4*). In higher eukaryotes it was reported to co-regulate transcription initiation and elongation (reviewed in (*5*)). Therefore, it is one of promising candidates, which might be involved in the coupling of transcription to splicing.

Another fundamental splicing-related question is how splicing machinery recognizes the sequences to be excised from pre-mRNA. On the sequential level, introns are defined only by three conserved motifs: the 5′ splice site and 3′ splice site on the respective ends of intron, and the sequence flanking the branch point adenosine. Recognition of 5′ splice site and branch point is based on the base pairing with the U1 and U2 snRNAs, respectively (*6-8*). Mechanism of 3′ splice site recognition is still not satisfactory understood. In the second part of this work we document that secondary structures downstream of branch point are required for the proper recognition of remote 3′ splice sites in *Saccharomyces cerevisiae*.

Aims of the study

I. Elucidate the role of protein Prp45 in pre-mRNA splicing in *Saccharomyces cerevisiae.*

II. Explain the mechanism of distant 3′ splice site recognition in *S. cerevisiae*.

Materials and methods

Most experiments were performed with the *S. cerevisiae* strain EGY48 (*MATα his3 trp1 ura3* $LexA_{op(x6)}$ – *LEU2*) and its *prp45*(1-169)::kanMX4 knock-in variant. Other strains were used where appropriate.

Synthetic lethal screen was based on random UV mutagenesis and plasmid shuffling technique (*9*).

Two hybrid interactions were monitored by the β-galactosidase colony lift filter assay in GAL4 AD/BD system (Clontech).

Levels of pre-mRNA and mRNA were quantified by quantitative RT-PCR technique on the LC480 instrument.

Splicing efficiencies were assayed by primer extension on *CUP1* fusion reporter genes expressed from plasmids under the control of the strong *TDH3* promoter. Outputs from the phosphorimager were optionally densitometrically quantified to calculate the first and the second step efficiency. For primer extension purpose a number of plasmids were constructed using standard gene engineering techniques.

For RNA structure modeling, multiple algorithms were employed, including RNAfold (*10*) and RNAshapes (*11*). Intron sequences were downloaded from public databases, mainly from SGD (*Saccharomyces* genome database).

Results

In the first part of study, we present an analysis of the prp45(1-169) allele. We performed a comprehensive screen to identify alleles, which are synthetically lethal with *prp45*(1-169).

To complement the SL screen results, several previously characterized temperature sensitive mutants of splicing factors were tested for the synthetic lethality with *prp45*(1-169). Altogether, we identified ten genes with a mutation synthetic lethal with *prp45*(1-169). Among them, two main groups of splicing factors occur: (i) NTC members (Clf1/Syf3, Syf1, Ntc20, Cef1) and (ii) second step factors (Prp22, Prp18, Slu7 and Prp17). Only one gene identified (*COF1*) does not encode a splicing factor. Notably, the SL mutation in *COF1* localizes to the intron. All genetic interaction identified are thus in agreement with the role of Prp45 in splicing.

The analysis of Cwc2-TAP pull-downs revealed significantly decreased partition of Prp22 in the spliceosomal complexes isolated from *prp45*(1-169) cells in comparison to the wild-type levels. Indeed, we confirmed previously reported two-hybrid interaction of Prp45 and Prp22 (*4*). However, we did not observe the Prp22 interaction with Prp45(1-169), while the interaction of Prp22 with the C-terminal fragment of Prp45 is positive.

Using a set of *ACT1-CUP1* based reporter genes, we documented that *prp45*(1-169) exacerbates splicing defects of the substrates with suboptimal 5´ss, BP, and 3´ss sequences, while the canonical *ACT1* intron remains unaffected. Primer extension

analysis of substrates containing other natural yeast introns revealed that Prp45 is required for efficient splicing of specific introns, represented here by *COF1* intron.

We showed that the ectopic expression of Prp45(119-379) fragment suppressed most *prp45*(1-169) related phenotypes, including the temperature sensitivity, splicing defects, and impaired Prp22 partition in spliceosomes. These experiments suggest that the proper Prp45 functioning in splicing can be achieved by the expression of two (overlapping) parts *in trans*.

In an effort to explore a conservancy of the Prp45 functioning in splicing, we excluded the possibility that Prp45 homologs from other organisms (*S. pombe, A. thaliana*, and humans) could complement the inviability of *prp45*-Δ cells. However, AtSKIP, the Prp45 homolog from the most distant species *A. thaliana*, was completely able to suppress the temperature sensitivity and partially also splicing defects of *prp45*(1-169) cells, while human SKIP had only an intermediate positive effect, and SNW1 from *S. pombe* improved significantly nor the growth neither the splicing.

In the second part, we showed that splicing of introns with remote 3´ss is facilitated by formation of a secondary structure downstream of BP. We identified a single-nucleotide G to A

substitution between the BP and 3'ss in *COF1* intron which virtually eliminated its splicing. Independent algorithms predicted the formation of a long stem-loop structure in this region. The G to A substitution disrupted the stability of the modeled structure. We then mutated C to U in a position corresponding to the complementary nucleotide in the predicted stem, which stabilized the stem-loop according to the prediction, and found that this mutation restored splicing to wild-type levels. We also randomized the sequence of the 10 nucleotide stem region such that the stability of the predicted structure remained the same and showed that the *COF1* splicing was not impaired. Thus, the secondary structure, rather than a sequence specific interaction, is critical for efficient splicing. We tested *COF1* introns with other modifications or deletions within the BP-3´ss region. We found cryptic 3´splice sites that are apparently masked by the secondary structure but that become active upon its destabilization.

In addition, we manipulated the sequence of *UBC13* intron, which also has a long BP-3´ss distance, and showed that the disordering of the predicted secondary structure results in use of BP-proximal cryptic 3´ss. When the base pairing (but not the original sequence) in the predicted stem was restored through a set of mutations, the wild-type splicing pattern was observed. Mutations which caused only mild decrease in the stability of the

predicted structure resulted in alternative 3´ss choice which was temperature dependent.

RNAfold modeling predicts similar stem-like structures in majority of long BP-3´ss distance introns in *S. cerevisiae*. We reason that secondary structures aid efficient recognition of remote 3´splice sites in *S. cerevisiae* introns, both by shortening the structural distance between BP and the 3´splice site and by masking proximal cryptic sites.

Discussion

Prp45 was implicated in splicing based on: (i) its co-purification with spliceosomal complexes (*12, 13*), (ii) its two hybrid interaction with splicing factors Prp46, Prp22 and Syf3, (iii) the observation that the cells with downregulated *PRP45* expression accumulate unspliced precursors (*4*), and (iv) its homology with human SNW/SKIP, which is a component of the active spliceosome (*14*) and affects alternative splicing of specific transcripts (*15*).

Our documentation of *prp45* genetic interactions with the second step factors and the Prp45-Prp22 two hybrid interaction suggest that Prp45 may be involved in later stages of splicing. We documented the functional interaction of Prp45 with Prp22, helicase involved in the second step and mRNA release.

Biochemical analysis revealed significantly decreased levels of Prp22 in spliceosomal complexes isolated from *prp45*(1-169) cells. The C-terminal part of Prp45 may be required for the Prp22 recruitment to spliceosome. This hypothesis is supported by the fact that the C-terminus of Prp45 is indispensable for the Prp45-Prp22 interaction. Alternatively, the reduced Prp22 signal in the Cwc2-TAP pull-down from *prp45*(1-169) cells may reflect shortened period for which Prp22 persists in spliceosome. Indeed, the *prp45*(1-169) dependent acceleration of the Prp22 performance would be in agreement with the decreased splicing efficiency of suboptimal substrates observed in *prp45*(1-169) cells. Such substrates are spliced slowly and as a consequence are preferentially discarded by helicase based proofreading mechanisms (*16*, *17*). Higher rate of the Prp22 mediated rearrangements would result in increased rejection of these suboptimal slow substrates.

The functional interaction of Prp45 with Prp22 and the splicing defects of the 3´ss mutated substrates in *prp45*(1-169) cell suggest an engagement of Prp45 in the second step of splicing. However, Prp45 enters into splicing process before the first transesterification. Impairment of the substrates with suboptimal 5´ss and BP in the *prp45*(1-169) background suggest also a first step related role of Prp45. Our recent unpublished data showing

the accumulation of unspliced pre-mRNA molecules in *prp45*(1-169) cells indicate that Prp45 affects also the first splicing reaction.

Data published here and also our additional preliminary observations clearly show differential behavior of intron containing transcripts in *prp45*(1-169) cells. We are currently trying to identify the pre-mRNA hallmark(s) that are responsible for the splicing defects resulting in decreased amounts of respective spliced mRNA in the *prp45*(1-169) mutant. Inefficient splicing is partially improved by expression of Prp45 homolog from *A. thaliana* (AtSKIP), and to some extent also by human SNW/SKIP and SNW1 from *S. pombe*. Hence, it seems that the functioning of SNW proteins in splicing is conserved across eukaryotes, although the sequential differences do not enable complete rescue of the phenotype.

We showed that formation of a secondary structure between the BP and 3´ss is indispensable for the proper 3´ss recognition and as a consequence for efficient splicing of *COF1* and *UBC13* introns in *Saccharomyces cerevisiae*. The same phenomenon was independently documented for other four introns with distant 3´ss (*18*). The secondary structures either bring remote 3´ss to the physical limits of a spliceosomal interaction range or shorten the

time required for the distant 3´ss recognition, thereby preventing from rejection by a kinetic proofreading mechanisms (*19*).

We showed that the sequence forming the stem loop structure in the *COF1* intron can be deleted without any detectable impact on splicing. It begs the question, why the sequences forming the structures with no apparent benefits for splicing reaction itself were not eliminated in the evolution. We speculate about the role in fine regulation of splicing efficiency, which can be crucial only under some specific conditions.

As secondary structures mediate 3´ss recognition in majority of introns with distant 3´ss in *S. cerevisiae*, it would be interesting to know, whether this mechanism is conserved in other species. All introns with distant 3´ss from *Saccharomycotina* species we tested are predicted to form stable stem-like structures downstream of BP, indicating that the structure based mechanism might be plausible in multiple yeasts. Surprisingly, in all human introns tested, the region downstream of BP is predicted to be completely unstructured. Thus, it strongly suggests that the secondary structure based mechanism for 3´ss recognition is not generally applicable in higher eukaryotes. Taken together, variability in distributions of distances between BP and 3´ss together with dissimilar (predicted) ability of dBP introns from unrelated species

to fold into stable secondary structure suggest the existence of multiple mechanisms for 3´ss recognition.

Conclusions

- We identified ten genes with mutation(s) synthetic lethal with *prp45*(1-169).
- The C-terminus of Prp45 is indispensable for the interaction of Prp45 with the Prp22 RNA helicase in two hybrid system.
- The *prp45*(1-169) allele exacerbates splicing defects of *ACT1-CUP1* substrates with suboptimal 5´ splice site, branch point, or 3´ splice site sequences, while the canonical *ACT1* intron is spliced with the wild-type efficiency.
- Prp45 differentially affects splicing of diverse introns. A severe splicing defect in *prp45*(1-169) cells was documented for *COF1* intron.
- The *prp45*(1-169) related defect phenotypes can be suppressed by the ectopic expression of the C-terminal portion of Prp45.
- The *prp45*(1-169) splicing defects can be partially suppressed by the expression of AtSKIP and SNW/SKIP, the *A. thaliana* and human Prp45 orthologs, respectively.

- Formation of a secondary structure between the branch point and the 3´ splice site in *COF1* and *UBC13* introns decreases an „effective“ distance between the branch point and the 3´ splice site and mask potential cryptic 3´ss. Both mechanisms are required for the proper 3´ss recognition and for the efficient splicing. Resembling secondary structures are predicted in most *S. cerevisiae* introns with remote 3´splice sites.
- The secondary structures in *COF1* and *UBC13* introns are conserved in all species from *Saccharomyces sensu stricto* genus.

References

1. G. Orphanides, D. Reinberg, *Cell* **108**, 439 (2002).
2. M. C. Wahl, C. L. Will, R. Luhrmann, *Cell* **136**, 701 (2009).
3. C. L. Will, R. Luhrmann, *Cold Spring Harb. Perspect. Biol.* **3**, (2011).
4. M. Albers, A. Diment, M. Muraru, C. S. Russell, J. D. Beggs, *RNA* **9**, 138 (2003).
5. P. Folk, F. Puta, M. Skruzny, *Cell Mol. Life Sci.* **61**, 629 (2004).
6. R. Parker, P. G. Siliciano, C. Guthrie, *Cell* **49**, 229 (1987).
7. M. Rosbash, B. Seraphin, *Trends Biochem. Sci.* **16**, 187 (1991).
8. J. Wu, J. L. Manley, *Genes Dev.* **3**, 1553 (1989).
9. L. Barbour, W. Xiao, *Methods Mol. Biol.* **313**, 161 (2006).
10. A. R. Gruber, R. Lorenz, S. H. Bernhart, R. Neubock, I. L. Hofacker, *Nucleic Acids Res.* **36**, W70 (2008).
11. P. Steffen, B. Voss, M. Rehmsmeier, J. Reeder, R. Giegerich, *Bioinformatics.* **22**, 500 (2006).
12. P. Fabrizio *et al.*, *Mol. Cell* **36**, 593 (2009).
13. M. D. Ohi *et al.*, *Mol. Cell Biol.* **22**, 2011 (2002).
14. S. Bessonov, M. Anokhina, C. L. Will, H. Urlaub, R. Luhrmann, *Nature* (2008).
15. V. Bres, N. Gomes, L. Pickle, K. A. Jones, *Genes Dev.* **19**, 1211 (2005).
16. R. M. Mayas, H. Maita, J. P. Staley, *Nat. Struct. Mol. Biol.* **13**, 482 (2006).
17. M. M. Konarska, J. Vilardell, C. C. Query, *Mol. Cell* **21**, 543 (2006).
18. M. Meyer, M. Plass, J. Perez-Valle, E. Eyras, J. Vilardell, *Mol. Cell* **43**, 1033 (2011).
19. C. C. Query, M. M. Konarska, *Nat. Struct. Mol. Biol.* **13**, 472 (2006).

Publications and CV

Publications related to the thesis:

Gahura O., Abrhámová K., Skružný M., Valentová A., Munzarová V., Folk P., Půta F.
Prp45 affects Prp22 partition in spliceosomal complexes and splicing efficiency of non-consensus substrates.
J Cell Biochem 2009, 106(1):139-151 (IF 3.381)

Gahura O., Hammann Ch., Valentová A., Půta F., Folk P.
Secondary structure is required for 3' splice site recognition in yeast.
Nucleic Acids Research, in press (IF 7.836)

Wang X., Wu F., Xie Q., Wang H., Wang Y., Yue Y., Gahura O., Ma S., Liu L., Cao Y., Jiao Y., Puta F., McClung C. R., Xu X., Ma L.
SKIP is a splicing factor linking alternative splicing and circadian clock in *Arabidopsis*
Plant Cell, manuscript submitted (IF 9.396)

Oral presentations:

Gahura O., Valentová A., Abrhámová K., Folk P., Půta F.
Prp45-Prp22 interaction is required for splicing of non-consensus substrates in Saccharomyces cerevisiae, RNA Club 2008, Prague, Czech Republic

Gahura O., Půta F., Folk P.
Splicing of *Saccharomyces cerevisiae* introns with long branch point 3' splice site distance depends on secondary structure

39th Annual Conference on Yeasts, 3-6 May 2011, Smolenice, Slovakia

Gahura O., Valentová A., Půta F., Folk P.
Secondary structure formation facilitates splicing of *Saccharomyces cerevisiae* introns with long branch point 3' splice site distance
25th International Conference on Yeast Genetics and Molecular Biology, 11-16 July 2011, Olsztyn, Poland; Yeast, Vol. 28, Issue S1, p. S139

Gahura O., Valentová A., Půta F., Folk P.
Intron secondary structure mediated 3´splice site recognition in *Saccharomyces cerevisiae*
Czech – Spanish Workshop on Yeast Physiology, Institute of Physiology, AS of Czech Republic, 19 – 20 September 2011, Prague, Czech Republic

Gahura O., Valentová A., Půta F., Folk P.
Intron secondary structure mediated 3´splice site recognition in yeast
RNA Club 2011, 11 November 2011, Prague, Czech Republic

Poster presentations (presenting author):

Gahura O., Valentova A., Abrhamova K., Folk P., Puta F. Prp45, the homolog of SNW1/SKIP, functionally interacts with the DEAH-box helicase Prp22 to affect splicing fidelity in *S. cerevisiae.*
Experimental Biology 2008, San Diego, USA, 5-9 April 2008; FASEB J. 2008 22: lb210.

Gahura, O, Valentova, A, Abrhamova, K, Folk P., Puta F.
Transcriptional co-regulator homolog Prp45 is required for splicing of non-consensual substrates in yeast.
33rd FEBS Congress, Athens, Greece, 28 June – 3 July 2008; FEBS Journal, Vol. 275, Supplement 1, p. 121.

Gahura O., Konirova J., Valentova A., Folk P., Puta F.
Intron-specific effects of Prp45, homolog of human transcriptional co-regulator SNW/SKIP, in *Saccharomyces cerevisiae*
EURASNET Interdisciplinary Focus Meeting, Frontiers in Structural Biology of RNAs and RNPs, 16-18 August, 2010, Poznan, Poland

Gahura O., Puta F., Folk P.
Long branch point 3' splice site distance requires secondary structure formation for efficient splicing in *Saccharomyces cerevisiae*
Second International EURASNET Conference on Alternative Splicing, 28 February – 3 March 2011, Granada, Spain