ANALYSIS OF YEAST PRP8 PROTEIN AND
ITS ROLE IN U5 SNRNP ASSEMBLY

KUM LOONG BOON

A THESIS PRESENTED FOR THE DEGREE OF DOCTOR OF
PHILOSOPHY

UNIVERSITY OF EDINBURGH
DECEMBER 2005
DECLARATION

I hereby declare that I alone composed this thesis and that the work presented here is my own, except where stated otherwise.

December 2005
Acknowledgements

Firstly, I am indebted to the kindness and generosity of Professor Ken Murray and the Darwin Trust of Edinburgh for funding me during my PhD. It was not possible to further my study in the absence of support.

Dear Jean, thanks for giving me the opportunity to work in your lab, your constant encouragements, and for your help and suggestions over the past three years. And the care that you have shown which is beyond the duty of a supervisor is most appreciated.

To Michael (Dr. Spiller), thanks for helping me with the microscopy work, writting of the thesis, and for being such a nice labmate in the past 3 years. Also, I would like to thank the past and present members of the Beggs, Tollervey, Hardwick labs, and friends for work and fun times.

I would like to acknowledge; Dr. Andrew Newman and Dr. Gretchen Edwalds-Gilbert for the fruitful collaborations, and again to Dr. Andrew Newman for help throughout my studies, Dr. Patrizia Fabrizio, Dr. Anita Hopper, and Dr. Yves Henry for yeast strains and reagents, Dr. Duncan Short for assistance with mass spectrometry analysis, Prof. David Tollervey and Dr. Christophe Dez for work and champagne.

To my late father, my family, and especially to my sister Jin-Moi, the love, the care, and everything that you giving me are beyond the phrase “thank you".
Abstract

Prp8 protein is a component of the nuclear pre-mRNA splicing machinery. Prp8p had previously been identified as a component of the U5 snRNP, and the U4/U6.U5 tri-snRNP of the spliceosome. The protein contacts the catalytic centre of the spliceosome, as detected by photochemical crosslinks of Prp8p with functionally important regions of the U5 and U6 snRNAs, and with the 5' splice site, 3' splice site, and branchpoint of pre-mRNA.

In this thesis I present further work on protein interactions between Prp8p and other splicing factors by using Prp8p partitioned strains. The Prp8p partitioned constructs allowed investigation of factors that associate with the N- or C-terminal regions of Prp8p by isolation of the sub-complexes associated with different Prp8 protein fragments. Thus, the U5 snRNP protein Snu114p associates with Prp8p in the region 437-770, whereas fragmenting Prp8p at residue 2173 destabilises its association with Aar2p.

The role of Prp8p in U5 snRNP assembly was investigated. Upon identification of the nuclear localisation signal of Prp8p, I further investigated the role of Prp8p in U5 snRNP assembly, and showed that Prp8p-Snu114p-Aar2p-U5 complex formation happened in the cytoplasm, and that formation of the Prp8p-Snu114p-Brr2p-U5 complex took place in the nucleus. Despite the fact that Prp8p remains nuclear localised, the association between Prp8p and Brr2p is disrupted by Prp8p C-terminal mutations that resemble human Prp8p mutations that result in retinitis pigmentosa type 13. In addition, I also identified that the importin, Kap95p, is needed for the nuclear import of Snu114p and snRNA nuclear import. Based on the results above, a model of the yeast U5 snRNP assembly pathway is proposed.

In this thesis, I also characterise a Prp8p co-purified protein, Spp382p. The characterisation of Spp382p showed that this protein binds to U2, U5, U6 snRNAs, and excised intron, and interacts with Prp43p directly. Genetic-depletion of Spp382p results in excised intron accumulation, suggesting that this protein is needed for the late-stage of splicing. In addition, in Spp382p-depleted extract the U5 snRNP protein, Prp8p, is associated with 8-fold more U2 snRNA. Also, glycerol gradient analysis of snRNPs showed reduced levels of free U5, and U6 snRNPs, but accumulation of U4/U6 di-snRNP, suggesting a defect in recycling snRNPs when compared to non-depleted cell extract. Interestingly, in the Spp382p-depleted extract, Prp43p was not able to pull down excised intron. Overall results suggest that Spp382p could be needed for disassembly of post-splicing complex and to mediate the association of Prp43p with excised intron.
Common Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>Amino acid(s)</td>
</tr>
<tr>
<td>Amp</td>
<td>Ampicillin</td>
</tr>
<tr>
<td>ARS</td>
<td>Autonomous replication sequence</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine 5'-triphosphate</td>
</tr>
<tr>
<td>bp</td>
<td>Base-pair</td>
</tr>
<tr>
<td>°C</td>
<td>Degree Celsius</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary DNA</td>
</tr>
<tr>
<td>CEN</td>
<td>Centromere element</td>
</tr>
<tr>
<td>cpm</td>
<td>Counts per minute</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DNase</td>
<td>Deoxyribonuclease</td>
</tr>
<tr>
<td>dNTP</td>
<td>Deoxynucleoside triphosphate</td>
</tr>
<tr>
<td>ECL</td>
<td>Enhanced Chemi-Luminescence</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescence in situ hybridisation</td>
</tr>
<tr>
<td>FTC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>g</td>
<td>Gram(s)</td>
</tr>
<tr>
<td>h</td>
<td>Hour(s)</td>
</tr>
<tr>
<td>HEPES</td>
<td>N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid</td>
</tr>
<tr>
<td>His</td>
<td>Histidine</td>
</tr>
<tr>
<td>Hph</td>
<td>Hygromycin</td>
</tr>
<tr>
<td>HRP</td>
<td>Horseradish Peroxidase</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>k</td>
<td>Kilo</td>
</tr>
<tr>
<td>kb</td>
<td>Kilobase</td>
</tr>
<tr>
<td>L</td>
<td>Litre(s)</td>
</tr>
<tr>
<td>Mb</td>
<td>Megabase</td>
</tr>
<tr>
<td>min</td>
<td>Minute(s)</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>nt</td>
<td>Nucleotide</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RNase</td>
<td>Ribonuclease</td>
</tr>
<tr>
<td>RNP</td>
<td>Ribonucleoprotein particle</td>
</tr>
<tr>
<td>rNTP</td>
<td>Ribonucleoside triphosphate</td>
</tr>
<tr>
<td>rpm</td>
<td>Revolutions per minute</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>s</td>
<td>Second(s)</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulphate</td>
</tr>
<tr>
<td>SGD</td>
<td>Saccharomyces Genome Database</td>
</tr>
<tr>
<td>snRNA</td>
<td>Small nuclear RNA</td>
</tr>
<tr>
<td>snRNP</td>
<td>Small nuclear ribonucleoprotein particle</td>
</tr>
<tr>
<td>ssDNA</td>
<td>Single stranded DNA</td>
</tr>
<tr>
<td>SV40</td>
<td>Simian virus 40</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>tRNA</td>
<td>Transfer RNA</td>
</tr>
<tr>
<td>V</td>
<td>Volt(s)</td>
</tr>
<tr>
<td>W</td>
<td>Watt(s)</td>
</tr>
<tr>
<td>w/v</td>
<td>Weight per unit volume</td>
</tr>
<tr>
<td>yePCR</td>
<td>Yeast colony PCR</td>
</tr>
</tbody>
</table>
# Table of Contents

## Chapter 1

### Introduction

1.1 Pre-mRNA splicing  
1.2 The spliceosome machinery  
1.3 Splicing complex regeneration  
1.4 Group II introns  
1.5 Exon definition  
  1.5.1 The Cis-acting Intron determinants  
  1.5.2 The 5' splice site (U2-dependent introns)  
  1.5.3 The Branchpoint  
  1.5.4 The 3' splice site and polypyrimidine tract (U2-dependent introns)  
  1.5.5 The U12 dependent pre-mRNA  
1.6 Alternative splicing  
1.7 Canonical Sm proteins  
1.8 Biogenesis of spliceosomal snRNPs  
1.9 Prp8p  
1.10 Retinitis pigmentosa  
1.11 U5 snRNP  
1.12 Links between pre-mRNA splicing and other mRNA processing  
1.13 The aim of the current studies

### Chapter 2

### Materials and Methods

2.1 General reagents  
  2.1.1 Bacterial and yeast growth media  
  2.1.2 Commonly used buffers  
  2.1.3 Oligonucleotides  
  2.1.4 Plasmid vectors and constructs  
  2.1.5 E. coli strains  
  2.1.6 *Saccharomyces cerevisiae* strains  
  2.1.7 Antisera  
  2.1.8 Markers  
2.2 Microbiological methods
Chapter 3

Prp8p Partitioned Reveals Protein-protein Interactions

3.1 Introduction 55
3.2 Partitioned Prp8p 55
3.3 Partition strain growth phenotypes 58
3.4 Requirement for 3 amino acids repeats and Mosaic End sequences for cell viability in the Prp8p aa436 partition strain 58
3.5 TAP-tag purification of Prp8p complexes 60
3.6 Co-immunoprecipitation of Snu114p and Aar2p with Prp8p 62
3.7 Discussion 66
Chapter 4

U5 snRNP assembly, from cytoplasm to nucleus

4.1 Introduction 68
4.2 Prp8p NLS identification 68
4.3 Effect of Prp8p delocalisation on localisation of other U5 components 72
4.4 Prp8p mediates Aar2p nuclear localisation 75
4.5 Retinitis pigmentosa type 13 in yeast 76
4.6 Prp8p, Aar2p, Brr2p and U5 snRNA associations 76
4.7 Yeast U snRNP importin identification 83
4.8 Discussion 88

Chapter 5

Prp8p co-purified protein, Spp382p, is required for late-stage of splicing

5.1 Introduction 91
5.2 Co-immunoprecipitation of snRNAs and spliceosome with Spp382p 93
5.3 Depletion of Spp382p 95
5.4 Spp382p distribution and association with Prp43 protein 98
5.5 Spp382p mediates Prp43p function in post-splicing complex 100
5.6 Nuclear localisation of Spp382p, and Prp43p in relation of Spp382p Expression level 101
5.7 Discussion 103

Chapter 6

Final Discussion and Future Work

6.1 Future work – Part one 106
6.2 Future work – Part two 107

References 108
Chapter 1
Introduction

1.1 Pre-mRNA splicing

In 1970s, it was discovered that RNA in the nucleus of vertebrate cells was unusually long, compared with the shorter mRNA that emerged in the cytoplasm. This nuclear RNA had both a cap structure at its 5' end and a polyadenosine [poly(A)] tract at its 3' end, just like the shorter cytoplasmic mRNA. Resolution of these paradoxes occurred with the discovery of split genes and RNA splicing. Non-coding sequences (introns) are normally removed from newly transcribed RNA and the regions present in the mature RNA (exons) are joined in a process termed RNA splicing (Berget et al., 1977; Chow et al., 1977). Four different classes of RNA splicing are currently known, exemplified by the splicing of 1) group I, 2) group II, 3) nuclear precursor messenger RNA (pre-mRNA), and 4) tRNA introns (Dix and Beggs, 2001).

Metazoan genes are estimated to average nearly 10 introns (Ares et al., 1999), which must be precisely removed to yield translatable mRNAs. The presence of multiple introns can allow the joining of different combinations of exons by alternative splicing and produce distinct mRNAs from identical pre-mRNAs, thereby increasing the informational capacity of the genome. *Saccharomyces cerevisiae* (budding yeast) is ideal as a model for pre-mRNA splicing analysis, as fewer than 250 of the more than 6200 annotated genes are known to have introns (Ares et al., 1999). However, the distribution of introns appears to be non-random; about 70% of the highly expressed genes that encode ribosomal proteins contain an intron (Mayes et al., 2000). Most budding yeast nuclear introns are fewer than 600 nucleotides, and are generally found towards the 5' end of the pre-mRNA, such that the first exon is small. However, there are exceptions, for example the *DBP2* gene contains an intron that is more than 1000 nucleotides in length and lies 1273 nucleotides from the initiating AUG (SGD). Only a few budding yeast nuclear genes are known to have more than one intron, including *TAD3, WMA9, SUS1, MATa1, DYN2, RPL7A, RPL7B*, and *RPS22B* (Mayes et al., 2000, SGD).

Introns are highly divergent in sequence. However, three short regions of consensus sequence have been identified that define introns and contribute chemically reactive nucleotides to the splicing reaction: the 5' splice site, branchpoint, and 3' splice site (Figure 1.1 & 1.2). Pre-mRNA splicing occurs via an ordered pathway involving two sequential
trans-esterification reactions (Figure 1.1). In the first trans-esterification step, the phosphodiester bond at the 5' splice site is cleaved as a result of nucleophilic attack by the 2' hydroxyl group of a conserved adenosine that lies in the intron towards the 3' splice site, and that is referred to as the branchpoint nucleotide. This yields two intermediates: exon 1 with a free 3' hydroxyl, and intron-exon2 in a branched-lariat structure, in which the 5' end of the intron is linked covalently to the branchpoint adenosine via a 2'-5'phosphodiester bond. In step 2, cleavage at the 3' splice site occurs as a result of nucleophilic attack by the free 3' hydroxyl of exon 1. The intron is released, and the exons are simultaneously joined via a 3'-5' phosphodiester bond, producing the two products of the splicing reaction: a mature messenger RNA (mRNA) and an excised intron-lariat (Figure 1.1).

1.2 The spliceosome machinery

Nuclear pre-mRNA splicing in yeast and metazoa is catalysed by a large ribonucleoprotein complex, the spliceosome. This is assembled onto each intron in an ordered, multi-step process from the snRNPs (small nuclear ribonucleoproteins) U1, U2, U4/U6, U5, as well as non-snRNP proteins (Will and Luhrmann, 2001; Jurica and Moore, 2003), the exact nature and number of which has been unclear. The snRNPs are trans-acting factors with both RNA and protein components, and are named according to the snRNA each contains. The snRNA-associated proteins fall into two groups. Seven core or Sm proteins (B, D1, D2, D3, E, F, and G as originally defined in metazoa) are common components of the U1, U2, U4, and U5 snRNPs, while each snRNP also contains snRNA-specific proteins.

This traditional view of spliceosome assembly involves the sequential interaction of snRNPs with the pre-mRNA in a highly ordered and stepwise manner (reviewed in Staley and Guthrie, 1998) (Figure 1.1). Assembly begins with the association of the U1 snRNP with 5’ splice site to form the Commitment Complex (CC). Neugebauer and Rosbash groups (Kotovic et al., 2003; Gornemann et al., 2005; Lacadie and Rosbash, 2005) have demonstrated that U1 recruitment to pre-mRNA appears to occur co-transcriptionally, requiring the presence of introns in the pre-mRNA. Subsequently, the U2 snRNP binds to the branchpoint sequence, via Watson-Crick base-pairing, forming the A complex. The two sequences are not completely complementary, resulting in the conserved adenosine residue at the branchpoint being unpaired and bulging out of the U2 snRNA-pre-mRNA duplex. This bulging of the adenosine residue allows it to attack the 5’splice site in the first step of splicing.
Figure 1.1. Schematic demonstrating the assembly of the spliceosome by ordered addition of snRNP particles onto the pre-mRNA. The U1 snRNP (brown) interacts with 5' splice site (5'ss) to form the Commitment Complex (CC). Following this, the U2 snRNP (grey) associates with an adenosine nucleotide in the intron (A), referred as branch-point (Bp) to form the A complex. Next, the U5.U4/U6 tri-snRNP associates with the complex to create the pre-spliceosome or Complex B. Complex C is then produced by the rearrangements within the spliceosome, including the dissociation of the U1 snRNP and the U4 snRNP (green) from the U6 (purple). The splicing chemistry then occurs. In the first step, the 2'hydroxyl group (OH) of the branch-point attacks the phosphodiester bond (p) at the 5' splice site. The bond breaks, and a new phosphodiester bond is formed (GpA). In the second step, the 3'OH group of the free 5' exon attacks the 3' splice site. The result is joined exons (the spliced mRNA) and excised intron. The spliceosome dissociates, releasing mature mRNA and the intron-lariat structure. snRNPs are then released and recycled back into the splicing pathway, while the lariat is debranched by the debranching enzyme (Dbr1p) and subsequently degraded (Adapted from Makarov et al., 2002 and Newman 2001).
Following a complex formation, the U4/U6.U5 triple snRNP associates with the complex, forming the spliceosomal B complex. The addition of the U4/U6.U5 triple snRNP to the spliceosome results in a number of RNA:RNA re-arrangements. The formation of the active spliceosome (C complex) is promoted by the Prp19p complex (NTC) (Chan et al., 2003). Association of NTC with B complex causes destabilisation of the base pairing between U4:U6 snRNAs and formation of RNA-RNA interactions between U2 and U6 (Umen and Guthrie, 1995). Also, U1 snRNA is displaced from binding to the 5' splice site by U5 and U6 snRNAs. Both the U1 and U4 snRNPs are thought to leave the spliceosome after these rearrangements, leaving the activated spliceosome based around the U2, U5 and U6 snRNPs.

Following pre-mRNA splicing, the spliced RNA is released from the spliceosome through the action of Prp22p (Schwer and Gross, 1998), and the spliceosome subsequently disassembles, releasing the lariat intron, which is de-branched by Dbr1p and degraded. Prp43p is a putative RNA helicase that is important for releasing the lariat intron and U2, U5, and U6 snRNPs (Arenas and Abelson 1997). It is thought that for another round of splicing to occur, the components have to be re-assembled.

An alternative process of spliceosome formation has been proposed by Stevens et al. (2002), who suggested that in yeast a penta-snRNP complex forms independently of pre-mRNA interaction. This proposal is supported by evidence using in vitro transcribed RNA, and penta-snRNPs purified from HeLa cell nuclear extracts, showing that U1 associated with the 5' splice site as part of a complex containing all five U snRNAs (Malca et al., 2003).

In vivo newly synthesised transcripts are usually spliced during transcription or immediately thereafter. This presumes that pre-mRNA splicing occurs exclusively in the cell nucleus. However, Denis et al. (2005) reported the presence of functional spliceosomes and signal-dependent pre-mRNA splicing in the cytoplasm of platelets. However, whether the pre-mRNA splicing in the cytoplasm is limited to platelet cells, or whether it might also occur in cell types that contain a nucleus, or if pre-assembly of part of the spliceosome components could happen in the cytoplasm before import into the nucleus, is unclear and further investigation is needed.
1.3 Splicing complex regeneration

Having performed the task of ligating exons and producing mRNA, the spliceosome must release the mRNA for export. Furthermore, the postcatalytic spliceosomal machinery must be reconfigured to allow a new round of splicing. The snRNP-bound lariat intron must be disassembled, allowing the excised lariat intron to be degraded and the snRNPs to be recycled. Release of mRNA, disassembly, and recycling all involve extensive RNA:RNA rearrangement. The base pairing of U6 to the 5' splice site, U2 to the branch point, and U5 to the exons must be severed. Moreover, the mutually exclusive pairings involving U2, U6, and U4 must be restored to their original conformations.

Prp22p and Prp43p were shown to be required for post-splicing complex disassembly (Schwer and Gross, 1998; Arenas and Abelson, 1997; Martin et al., 2002). Prp43p has been shown to be required for disassembly of the spliceosome and release of spliced intron (Martin et al., 2002) by analysing dominant negative Prp43p mutants. Prp22p is needed for disassembly of the spliceosome, and release of excised intron and spliced mRNA (Schwer and Gross, 1998; Wagner et al., 1998). Both of these factors belong to the ATP-dependent DExH-box RNA helicase family.

Prp24p, the Lsm proteins, and Aar2p were also shown to promote regeneration of pre-mRNA splicing activity. U4 snRNA is unwound upon the activation of the spliceosome, so must again be base-paired with U6 snRNA for the next round of splicing. Prp24p and the Lsm proteins (Lsm6p and Lsm7p) were shown to be required for re-annealing of U4 and U6 snRNAs (Raghunathan and Guthrie, 1998; Verdone et al., 2004). When Prp24p, Lsm6p, or Lsm7p was absent, unpaired U4 and U6 snRNPs accumulated; with time, splicing became inhibited.

In contrast to Prp24p and the Lsm proteins, Aar2p is associated exclusively with the simple form of the U5 snRNP (16S), which is required for in vivo splicing and later rounds of in vitro splicing (Gottschalk et al., 2001). Investigation of the Aar2-U5 snRNP by electron microscopy revealed that, despite its comparatively simple protein composition, the yeast Aar2p-U5 snRNP appears structurally similar to the human 20S U5 snRNP, and thus may be a core particle from which mature U5 snRNP and subsequently the U4/U6.U5 tri-snRNP are assembled (Gottschalk et al., 2001).
1.4 Group II introns

Group II introns have only been identified in chloroplasts, fungal mitochondria and in some eubacterial genomes. Like spliceosomal introns, group II intron splice via two sequential transesterification reactions that yield ligated exons and an excised lariat-intron. Self-splicing group II introns and the splicesome share many conserved and functionally important structural features (Villa et al., 2002). Strikingly, the catalytically important group II intron domain 5 (D5) (Figure 1.2A) was found to be capable of replacing the intramolecular stem-loop (ISL) (Figure 1.2B) of U6/U6atac in an in vivo functional assay (Shukla and Padgett, 2002). In addition, the overall similarity between group II intron splicing and the splicesome in their catalytic cores (Valadkhan and Manley, 2001), and evidence for catalytic Mg$^{2+}$ ions bound to RNA moieties (Sontheimer et al., 1997; Gordon and Piccirilli, 2001; Villa et al., 2002) in both splicing machineries suggest a link between these 2 splicing processes.

However, unlike pre-mRNA splicing, group II intron splicing reactions do not require a complex of splicing components. Although some group II introns self-splice in vitro, this reaction generally requires nonphysiological conditions, and in vivo, proteins (e.g. CRS1p; Ostersetzer et al., 2004) are required for intronic RNA to fold into a catalytically active structure.

Figure 1.2. Structural parallels between A) group II introns and B) spliceosomal snRNAs. Homologous structures in the two systems are highlighted in the same colour. Location of the AGC triad, branch site adenosine and Mg$^{2+}$ binding site are indicated in both systems. Domain V and U6 ISL are in red, and the exon-binding sequence (EBS1, 2, and 3) and U5 are in orange. Exons are shown in magenta and branch site regions in the two systems are in blue and green. (Figure is taken from Valadkhan 2005).
1.5 Exon definition

The removal of introns in higher eukaryotes requires the spliceosome to accurately recognise the 3' and 5' splice sites. The basic exon definition model proposes that in pre-mRNAs with large introns, the splicing machinery searches for a pair of closely spaced splice sites in an exonic polarity. When such a pair is encountered, the exon is defined by the binding of U1 to the 5' splice site, as well as associated 3' splice site-recognising factor (U2AF and SC35 in human). Such an exonic perspective of splice site recognition has been termed "exon-definition". Interestingly, SR proteins have not yet been found in budding yeast. The absence of SR proteins in budding yeast may not be surprising as this organism's introns are small, such that they might not use exon definition (Berget 1995).

1.5.1 The Cis-acting Intron determinants

Nuclear pre-mRNA introns lack conserved sequences, except for minimal elements at the splice site and the branchpoint, and between the branchpoint and the 3' splice site (Figure 1.3). This lack of conservation is remarkable in view of the huge size of some metazoan introns, and presents these organisms with the problem of ensuring splicing fidelity. Intron sequence determinants have been identified through sequence comparisons of both yeast and mammalian introns (Mount, 1982), and their contribution to function was determined by mutation.

1.5.2 The 5' splice site (U2-dependent introns)

The general consensus of the 5' splice site is AG/GURAGU (where / denotes the cleavage site; Mount, 1982). In yeast, the 5' splice site is more highly conserved (consensus R/GUAUGU). This sequence is required for splicing (Green, 1986). Mutations within this sequence generally abolish splicing in yeast (Vijayraghavan et al., 1986; Fouser and Friesen, 1986), or lead to alternative splice sites in mammalian systems (Aebi et al., 1986). In yeast, mutation of the invariant GU nucleotides often prevents lariat formation. As in mammals, this leads to the accumulation of lariat intermediates (Newman et al., 1985). Instances are known of point mutations in 5' splice sites (or 3' splice sites) in introns of genes such as
retinoblastoma, and cystic fibrosis trans-membrane conductance regulator, causing human genetic disease due to exon-skipping or cryptic site activation.

![Diagram](https://example.com/diagram.png)

**Figure 1.3.** Schematic representation of conserved sequence within the introns of *S. cerevisiae* and mammals. The position of the 5' splice site, branch point and the 3' splice site are indicated, where Y represents conserved pyrimidines and R represents conserved purine residues. The conserved adenosine residue is underlined. The conserved sequence of mammalian U12-type introns is obtained from Will and Lührmann, 2005.

### 1.5.3 The Branchpoint

Yeast introns contain the sequence UACUAAC (the branchpoint sequence), which is usually located 20-60 nucleotides upstream of the 3' splice site, and which is required for splicing (Pikielny et al., 1983). The 3'-most adenosine in this sequence (Figure 1.3, shown in bold) provides the nucleophile in step I and is the site of the lariat branch. In budding yeast, the sequence is almost completely conserved. Mutations within this sequence generally abolish splicing, although as with the 5' splice site, some mutations allow step I but not step II, resulting in the accumulation of lariat intermediates (Jacquier and Rosbash, 1986; Vijayraghavan et al., 1986). That mutations in a *cis*-acting sequence have this phenotype is suggestive of a proofreading mechanism, operating to ensure high fidelity in splicing.
Mammalian introns contain a similar, although less strictly conserved branchpoint, of consensus UNCURAC. Compared to yeast, mutational analysis of mammalian introns suggests a far less stringent requirement for a sequence in mammalian introns conforming to this consensus. When the β-globin branchpoint was deleted, multiple crytic sites were activated, although lariat formation always occurred at an adenosine, located 22-37 nucleotides upstream of the 3' splice site (Ruskin et al., 1985). Competition between duplicated branchpoint sequences in introns, which otherwise lack cryptic sites, shows that base substitutions or deletions within this sequence abolish splicing in vitro, and promote cis-competition. Therefore, the mammalian branchpoint makes a significant contribution to splicing efficiency (Reed and Maniatis, 1988).

1.5.4 The 3' splice site and polypyrimitidine tract (U2-dependent introns)

Both yeast and mammalian introns end in the dinucleotide AG. Mammalian introns have in addition, a preceding stretch of pyrimidine-rich sequence (the polypyrimitidine tract), and the splice site consensus is YnYAG (Y: pyrimidine bases) (Mount et al., 1983). The sequences around the yeast 3' splice site are not well conserved beyond the AG, except that in many cases the 3' splice site is preceded by pyrimidines. The distance between the branchpoint and the 3' splice site may be vary, but is usually 20-40 nucleotides.

In mammalian introns, mutation of the polypyrimitidine tract blocks the early steps of spliceosome assembly (Reed and Maniatis, 1985). Reed (1989) demonstrated that efficient lariat formation only occurred when the branchpoint was directly adjacent to a polypyrimitidine tract. Introns with a short polypyrimitidine tract (14 nucleotides) also required the AG dinucleotide for step I, whereas those with a long polypyrimitidine tract (26 nucleotides) underwent step I in the absence of an AG dinucleotide. In the latter type of intron, where the branch site is specified independently of the 3' splice site, there has been a suggestion that a scanning mechanism searches for the first AG downstream of the polypyrimitidine tract, which is used as the 3' splice site (Smith et al., 1989). Consistent with this, stable secondary structures inserted after the polypyrimitidine tract block step II, and when the AG is deleted, the next AG downstream is used as the 3' splice site (Smith et al., 1989).

In yeast introns, the branch site is specified primarily by the UACUAAC sequence, independent of pyrimidine-richness in the 3' region. A model substrate, truncated immediately after the branchpoint, underwent step I when ligated via its 3' end to a
homopolymeric oligoribonucleotide of any sequence. The only requirement for lariat formation was intron length (at least 29 nucleotides following the branchpoint; Rymond et al., 1987). The polypyrimidine tract, though not strictly required, greatly enhances the efficient usage of an adjacent AG nucleotide in an alternative 3’ splice site competition assay (Patterson and Guthrie, 1991).

In summary, the view is of tripartite cis-acting elements in the 3’ region of the intron. The sequences are recognised co-ordinately during early spliceosome assembly. Yeast and mammalian systems differ in that they have stringent requirements respectively for the branchpoint and a polypyrimidine tract, but the function of elements in the 3’ region is broadly similar. In both systems, step I can occur in the absence of a functional 3’ splice site.

1.5.5 U12 dependent splicing

A rare class of introns, called AT-AC or U12-dependent introns, was identified a decade ago (Jackson 1991). The first set of U12-type introns identified contained the dinucleotides AT and AC (genomic DNA sequences) at their 5’ and 3’ ends, respectively, and for this reason they were originally referred to as ATAC introns. Subsequent studies, however, revealed that some U2-type introns also end in AT-AC and conversely that a large number of U12-type introns have dinucleotides GT and AG at their 5’ and 3’ ends, respectively (Dietrich et al., 1997; Wu and Krainer, 1997) (Figure 1.3). In comparison to U2-type introns, consensus sequences delineating the U12-type 5’ splice site -A(or G)UAUCCUUU- and branch site -UCCUUAACU- are longer and more tightly constrained. The U12-type 3’ splice site is typically denoted by YAC/ or YAG/. However, in contrast to U2-type introns, there is no strict 3’ splice site sequence requirement for U12-type introns; various other dinucleotides can serve as U12-type 3’ splice sites (Dietrich et al., 2001; Levine and Durbin, 2001; Hastings et al., 2005). In addition, the average distance between the U12-type branch site and 3’ splice site (12-15 nucleotides) appears to be shorter and more highly constrained than that observed with U2-type introns (18-40 nucleotides) (Levine and Durbin, 2001) and, similar to U2-type introns present in yeast, U12-type introns lack a polypyrimidine tract.
1.6 Alternative splicing

There are probably more than 100,000 different proteins expressed at any one time in a human cell, but the report of only 32,000 human genes (Venter et al., 2001) thus came as a surprise. This basic disparity indicated that the number of human expressed-sequence (mRNA) forms was much higher than the number of genes, suggesting a major role for alternative splicing in the production of complexity. Alternative splicing is observed among pre-mRNAs from over a third of protein-coding genes. This yields different mRNAs from one gene by altering one or all of the following: 1) the transcription start site, thus modifying the 5' end of the RNA, 2) the site of cleavage and polyadenylation, thus altering the 3' end of the transcript, 3) the definition of exons, providing for a different assortment of the coding information (Goldstrohm et al., 2001). Alternative splicing can be divided into four general categories. 1) The choice to remove or not remove an intron (Figure 1.4A). 2) The alternative use of 5' splice sites that will change the length of an exon (Figure 1.4B). 3) The alternative use of 3' splice sites, which will also change the length of an exon by extending the 3' border of the exon (Figure 1.4C). 4) The choice between exon inclusion and exon skipping (Figure 1.4D).

![Figure 1.4](image-url)

**Figure 1.4. Types of alternative splicing.** A) The choice of an intron remove or not remove B) alternative 5' splice site. C) Alternative 3' splice site. D) Exon inclusion or exclusion.
1.7 Canonical Sm proteins

In both yeast and metazoa, each snRNP particle (except for U6) is composed of an U snRNA, a set of specific proteins associated with each particular snRNP, and a set of Sm or core proteins shared by all snRNPs. Sm proteins were named after Stephanie Smith, the first patient in which the systemic lupus erythematosus-associated anti-Sm autoimmune antibodies were identified (reviewed in Khusial et al., 2005). The core protein is composed of Sm proteins B, B' (in mammals), D1, D2, D3, E, F, and G, which assemble around the Sm site of the snRNA. However, in mammalian cells, SmB is replaced by an alternatively spliced variant SmB' or by a different gene product, the SmN protein. The B and B' proteins are encoded by the same gene, and are produced by alternative splicing (Chu and Elkon, 1991). In yeast, due to the lack of alternative splicing, there is no SmB' protein. The protein SmN is structurally similar to the SmB protein, however, is only found in neural tissue (Sharpe et al., 1990).

snRNP Sm proteins share two evolutionarily-conserved sequence motifs which are involved in Sm protein-protein interactions, suggesting that the Sm proteins may have arisen from a single common ancestor (Hermann et al., 1995). These two regions were termed the Sm motifs; Sm motif 1 (32 amino acids long) and motif 2 (14 amino acids long). Sm motifs are essential for complex formation between B' and D3 in vitro (Hermann et al., 1995). The Sm proteins bind to the Sm site on snRNAs, which is a single-stranded structural motif consisting of the sequence RA(U)₃₆GR (where R is a purine base), and is normally found between double hairpin loops (Branlant et al., 1982). Sm protein interaction with this sequence has proven to be resistant to mutation, with several mutations failing to disrupt the binding in vitro (Jones and Guthrie, 1990).

The existence of a seven-member complex binding a single-stranded RNA has suggested the possibility that the Sm proteins might assemble in a ring, and the RNA might pass through this ring. Structural studies have proven that an Sm ring does form, but the nature of the RNA interaction with the ring remains unclear. Initial electron microscopy showed that the E-F-G sub-complex appeared as a ring, which was shown biochemically to contain two copies of the E, F, and G proteins (Plessel et al., 1997).

The D1-D2 and the B-D3 sub-complexes have been crystallised, allowing the structure of these sub-complexes to be solved to atomic resolution (Kambach et al., 1999) (Figure 1.5A). Each Sm monomer appears to consist of a highly conserved fold, forming a barrel-type structure consisting of an N-terminal α-helix followed by a strongly bent five-stranded β-sheet. This allowed the modelling of the structure of the complete Sm complex,
which suggested that the Sm proteins associate to form a seven-member closed ring structure with a positively charged central hole. This structure allows the possibility for single-stranded, positive-charged RNA to pass through the central hole. Although this has not been demonstrated, Urlaub et al. (2001) have shown that RNA contacts the inner surface of the Sm ring.

Figure 1.5. (A) Structural models of the Sm heptameric ring. The left-hand side (a) represents the ribbon structure, with each subunit labelled. The surface representation is shown on the right (b) with electrostatic potential (blue, positive; red, negative). The two orientations are identical. Figure taken from Kambach et al. (1999). (B) Schematic representation of the assembly of the Sm protein core onto snRNAs in mammalian cells. Firstly, D1-D2 and E-F-G complexes are formed, which assemble into a pentameric complex in the absence of RNA, before binding to an U snRNA Sm site to form the subcore. Finally, a B-D3 dimer binds to the subcore, forming a proposed ring structure around the Sm site (Raker et al., 1996).
The demonstration that yeast snRNPs are immunoprecipitated by anti-Sm antibodies (Tollervey and Mattaj, 1987), and the observation that yeast snRNAs are able to assemble with Xenopus Sm proteins upon microinjection into oocytes suggested that yeast snRNPs contain a set of proteins homologous to the metazoan Sm proteins (Riedel et al., 1987). Biochemical studies indicated that the different yeast snRNPs contain polypeptides with sizes similar to those human core proteins. Some of these proteins (SmD1, SmD3, and SmE) were identified in the yeast genome sequence by homology to their human counterparts (Rymond 1993; Roy et al., 1995; Bordonne and Tarassov, 1996). Scanning the yeast genomic database revealed the presence of several putative Sm proteins, some of which turned out to represent canonical snRNP core proteins (Seraphin 1995).

In higher eukaryotes, binding of snRNA to the Sm complex is required for the import of both RNA and proteins (reviewed in Kiss 2004). In yeast, the situation is less clear, and the Sm proteins may be imported separately, with the snRNA remaining nuclear and interacting with the Sm complex in the nucleus. Wherever the location of snRNP formation, the C-terminal tails of the Sm proteins are important for their nuclear localisation in both yeast and higher eukaryotes (Bordonne 2000; Girard et al., 2004). Initial investigations were carried out in S. cerevisiae (Bordonne 2000), and these experiments showed that the basic C-terminal tails of SmB1p and SmD1p appear to be essential for the nuclear import of the yeast Sm complex. Nuclear import of proteins in yeast requires Kap95p (Importin-β in humans) (Lovine et al., 1995) and Srp1p (Importin-α in humans) (Enenkel et al., 1995).

Srp1p binds nuclear localisation signals in newly synthesised proteins in the cytoplasm, and then dimerises with Kap95p to enable nuclear import of the NLS-containing protein. Classical nuclear localisation sequences (NLS) contain clusters of basic residues (reviewed in Jans et al., 2000). Although SmB1p, SmD1p, and SmD3p do not contain a classical NLS, the C-termini of all three are rich in basic amino acids. Fusion of the C-terminal tails of SmB1 and D1 to GFP resulted in localisation of the fusion protein to the nucleus, but SmD3p does not (in yeast). Examination of Sm protein NLS sequences in human cells were performed in a similar way. In contrast to yeast, human SmD1p and SmD3p, but not SmB1p, possess NLS properties (Girard et al., 2004).
Assembly of the metazoan U snRNPs is a multi-step process, following an ordered pathway. With the exception of U6, the major spliceosomal snRNAs (U1, U2, U4, and U5) are synthesised by RNA polymerase II. In metazoa, after the transcription of snRNAs by RNA polymerase II, the precursor snRNAs, with a 5'-terminal 7-monomethylguanosine (m\(^7\)G) cap structure and extra nucleotides at the 3' end, are assembled into a large export complex that includes the m\(^7\)G-cap-binding complex (CBC), the phosphorylated adaptor for RNA export (PHAX) and the CRM1/RanGTP complex. In its phosphorylated form, PHAX functions by mediating the interaction between the CBC/U snRNA complex and the CRM1/RanGTP complex, allowing export from the nucleus. After export to the cytoplasm through the nuclear pore complex (NPC), PHAX becomes dephosphorylated, owing to GTP hydrolysis by Ran (Ohno et al., 2000). Which causing the complex to dissociate, and allowing PHAX to re-enter the nucleus via two NLS signals. This ensuring the directionality of U snRNA export, and allowing PHAX to be recycled (Segref et al., 2001; reviewed in Kiss 2004; reviewed in Will and Luhrmann, 2001) (Figure 1.6).

In the cytoplasm, the survival of motor neurons (SMN) protein complex facilitates the association of pre-snRNA with a heteroheptameric ring of seven Sm proteins to form the snRNP core complex on the Sm site of the snRNA. The association of Sm proteins with snRNAs occurs in a stepwise manner. First, the D1-D2 and E-F-G complexes are formed, which assemble into a pentameric complex before binding to the U snRNA, forming the subcore. Then a B-D3 dimer binds to the subcore, forming a proposed heptameric ring structure around the Sm site (Raker et al., 1996) (Figure 1.5B). This is followed by the hypermethylation of the snRNA m\(^7\)G cap by a methylase to form a methyl\(^{2,27}\)-guanosine cap structure, and 3' end maturation of the snRNAs (Seipelt et al., 1999; reviewed in Kiss 2004). This binding also generates a bipartite nuclear localisation signal (NLS), composed of the Sm core complex and the snRNA cap structure, that permits the import of snRNPs to the nucleus.

Competition experiments indicate that snRNPs are imported by specific receptors, and that U snRNP import is mediated by importin-\(\beta\), which functions in this process without the NLS-specific importin-\(\alpha\) receptor. In metazoa, snurportin1 enhances the m\(^3\)G-cap-dependent nuclear import of U snRNPs. This importin \(\alpha\)-like adapter recognises only the m\(^3\)G cap but not the Sm core NLS, indicating that at least two distinct import receptors interact with the snRNP bipartite NLS (Palacios et al., 1997).
Figure 1.6. Assembly of U snRNPs in higher eukaryotes. Newly transcribed U snRNA acquires a methyl guanosine (m’G) cap in the nucleus, to which the cap binding complex (CBC) binds. This then associates with phosphorylated PHAX which in turn binds to CRM1 (Xpo1 in budding yeast) which associates with Ran-GTP. This complex is exported to the cytoplasm where complex dissociates, owing to PHAX desphosphorylation and GTP hydrolysis by Ran. Assisted by SMN, which binds the snRNA after nuclear export, Sm proteins then assemble around the Sm site of the U snRNA, which leads to hypermethylation of the m’G cap, forming an m^{2,7}G cap. This enables immature snRNP re-import into the nucleus via snurportin1 and importin-β. Once there, the U snRNP associates with snRNP-specific proteins, probably in Cajal bodies, forming a mature U snRNP complex (Figure is taken from Kiss, 2004).
However, the m$_3$G cap structure plays a differential role in the nuclear import of distinct snRNAs in Xenopus oocytes: it is an essential signal component for U1 and U2 snRNAs, whereas U4 and U5 snRNAs can enter the nucleus without being hypermethylated (Fischer et al., 1991). In yeast, tri-methylGuanosine Synthase (TGS1), which is essential for hypermethylation of the m$_7$G caps of both snRNAs and snoRNAs is located in the nucleolus. Surprisingly, this gene is not essential.

Deletion of TGS1 confers a cold-sensitive phenotype and accumulation of U1 in the nucleolus, suggesting that the hypermethylation of the m$_7$G cap structure takes place in the nucleolus (Mouaiikel et al., 2002). The assembly of the Sm proteins on the Sm sites is a prerequisite for cap hypermethylation, since snRNA mutants lacking the Sm site (Mattaj 1986), or genetic depletion of yeast SmD1 and SmD3 proteins, leads to inefficient recognition of the snRNAs by anti-TMG antibodies (Rymond, 1993; Roy et al., 1995).

After re-entering the nucleoplasm, the import proteins dissociate and the newly imported snRNPs transiently accumulate in Cajal bodies (Sleeman and Lamond, 1999). In the Cajal bodies, the U1, U2, U4, and U5 snRNAs undergo site-specific pseudouridylation (Ψ) and 2’-O-methylation (m) directed by small Cajal body RNAs (scaRNAs), which specifically accumulate in Cajal bodies (Jady et al., 2003). In addition to common Sm core proteins, the mature U1, U2, U4, and U5 snRNPs also contain snRNA-specific RNP proteins. Since the snRNP-specific proteins appear to concentrate in Cajal bodies, they are believed to associate with core snRNPs in this organelle. Finally, mature spliceosomal snRNPs accumulate in the interchromatin region in structures called splicing speckles. Usually, speckles are located close to actively transcribed genes. They probably provide storage sites for snRNPs or function in snRNP recycling (Lamond and Spector, 2003).

In contrast to the Pol-II-specific Sm snRNAs, biogenesis of the Pol-III-transcribed U6 snRNA is confined to the nucleus. Before accumulating in the nucleoplasmic splicing speckles, the maturing U6 snRNA visits the nucleolus and Cajal body. Synthesis of U6 snRNA terminates with a short U stretch that serves as a transcription termination signal for Pol III. The 3’-terminal U stretch of the newly synthesised U6 transiently associates with the La protein, which provides stability for the nascent RNA and facilitates snRNP assembly (Wolin and Cedervall, 2002). Later, the La protein is replaced by the doughnut-like heteromer of seven Sm-like proteins (Lsm2, Lsm3, Lsm4, Lsm5, Lsm6, Lsm7, and Lsm8) (Achsel et al., 1999; Mayes et al., 1999).

In contrast to other U snRNAs, U6 RNA is not thought to leave the nucleus (Vankan et al., 1990; Boelens et al., 1995; Pante et al., 1997), apart from one study which detected U6 snRNA transiently in the cytoplasm of mouse fibroblasts (Fury and Zieve, 1996), and a
heterokaryon assay in yeast that found that all small nuclear RNAs tested passed through the cytoplasm (Olson and Siliciano, 2003). Wheat germ agglutinin is used to inhibit nuclear import of proteins that contain a classical NLS. In *Xenopus* oocytes, U6 RNA injected into the cytoplasm is imported into the nucleus, and this import is inhibited by wheat germ agglutinin, suggesting that U6 import occurs by the same pathway as protein import (Fischer et al., 1991; Michaud and Goldfarb, 1992).

During its maturation, the U6 snRNA also undergoes site-specific pseudouridylation and 2'-O-methylation. However, in contrast to Sm snRNAs, which are modified by scaRNPs in the Cajal body, 2'-O-methylation and probably also pseudouridylation of U6 snRNA is directed by snoRNAs, which reside in the nucleolus. This indicates that the U6 cycles through the nucleolus to undergo snoRNA-mediated modifications (Lange and Gerbi, 2000). The mature U6 and U4 snRNPs associate via RNA:RNA interaction to form the U4/U6 di-snRNP, that also binds the U5 snRNP to form the U4/U6.U5 tri-snRNP. The recent finding that p110 (Prp24p in budding yeast), a protein factor facilitating U4/U6 di-snRNP formation, is localised to the Cajal body, suggests that this process may take place in this organelle (Stanek et al., 2003).

Recent evidence has provided an insight into the formation of the tri-snRNP in mammalian cells (Schaffert et al., 2004). RNAi-based depletion of the human homologues of Prp3lp and the U5 snRNP protein Prp6p resulted in a decrease in tri-snRNP levels and accumulation of U4/U6 di-snRNP and free U5 snRNP. The di-snRNP RNAs accumulated in Cajal bodies, while free U5 snRNP accumulated in the speckles, suggesting that U5 transport to Cajal bodies is important in tri-snRNP formation.

The finding of Olson and Siliciano (2003) of U6 shuttling in yeast in heterokaryon assays contradicts previous evidence that the U6 snRNA is not exported from the nucleus under normal conditions (Boelens et al., 1995), as well as the recent suggestion that *S. cerevisiae* snRNAs might not be exported (Mouaike et al., 2002). Because of this, and because the RNAs appeared to transfer simultaneously, it was suggested that the shuttling observed in heterokaryon assays might be due to a global leaking of RNA. However, reduced transfer of U1 snRNA in a leptomycin-treated *cml1-lmb* strain, and the fact that all RNAs not only appeared outside the nucleus, but were actively taken up by the second nucleus suggests that a mechanism exists for export and re-import of snRNAs in *S. cerevisiae* (Olson and Siliciano, 2003).
1.9 Prp8p

Lee Hartwell and co-workers first identified RNA8 in 1967, through a genetic screen to isolate temperature-sensitive (ts) mutations in *S. cerevisiae* (review in Grainger and Beggs, 2005). RNA8 was later renamed PRP8 (PRP: pre-mRNA processing, Vijayraghavan et al., 1989). Temperature sensitive prp8-1 cells accumulated RPL59 (ribosomal protein 59) pre-mRNA (Larkin and Woolford, 1983) and were also defective in splicing in vitro of ACT1 pre-mRNA (Lustig et al., 1986), identifying Prp8p as an essential pre-mRNA splicing factor.

Yeast PRP8 was cloned by complementation of the heat-sensitive growth defect of the prp8-1 mutation (Jackson et al., 1988). The 7.4-kb transcript was detected at approximately one copy per yeast cell, with a product size of 280 kDa (SGD). Prp8p exhibits a high degree of conservation throughout all eukaryotes, with 61% identity overall between yeast and humans in its 2413 amino acids sequence (Hodges et al., 1995; Luo et al., 1999, Grainger and Beggs, 2005). Twenty-eight full-length sequences of PRP8 from 26 eukaryotic organisms are currently available (Grainger and Beggs, 2005). In plants, *Arabidopsis thaliana* has two copies of PRP8, one on chromosome 1, and the other on chromosome 4, that encode proteins with 92% identity (Grainger and Beggs, 2005). *Nicotiana tabacum* L. (tobacco) Prp8p (>200 kDa) was also detected by Western blot using Prp8p antibodies raised against yPrp8p (Kulesza et al., 1993). In humans Prp8p is encoded by a single gene with 42 exons (McKie et al., 2001). Presumably, human Prp8p’s function in pre-mRNA splicing is required in all tissues, and its expression is notably higher in skeletal and cardiac muscle (Luo et al., 1999). However, the mouse Prp8p is detected most strongly in mouse testis and ovary (Takahashi et al., 2001).

The use of Prp8p antibodies (Beggs and co-workers) has further identified its role in pre-mRNA splicing (Jackson et al., 1988; Anderson et al., 1989; Whittaker et al., 1990; Brown and Beggs, 1992). Immunodepletion of Prp8p with 8.1 antibodies (against amino acids 682-1122) from yeast extracts caused loss of splicing activity, and this antibody precipitated U5 snRNP and disrupted the U5.U4/U6 tri-snRNPs (Jackson et al., 1988; Lossky et al., 1987). In contrast, incubation of yeast splicing extracts with 8.4 antibodies (against aa1922-2410) or 8.6 antibodies (against aa2-35) (Brown and Beggs, 1992) resulted in co-immunoprecipitation of U4, U5 and U6 snRNAs, and the following spliceosomal complexes: pre-mRNA, excised-intron-exon2, and excised-intron. These data indicate that Prp8p is present in the spliceosome during the first and second steps of pre-mRNA splicing, and in a post-splicing complex (Jackson et al., 1988; Whittaker et al., 1990, Brown and Beggs, 1992). Polyclonal antibodies raised against the C-terminus of human Prp8p (70R and
K05), immunoprecipitated both the U2 and U12 spliceosomes, indicating that Prp8p is present in both major and minor spliceosomes (Hinz et al., 1996; Luo et al., 1999).

Budding yeast Prp8p N-termini contain proline-rich regions (Brown 1995, Grainger and Beggs, 2005). By yeast 2-hybrid analysis, only the Snu40/yLin1 protein (a non-essential U5 snRNP protein) (Giaever et al., 2002, SGD) was shown to interact with the N-terminus of yeast Prp8p (Bialkowski and Kurlandzka, 2002). The yeast strain with Prp8p proline-rich region deleted is viable, and shows no temperature-sensitive phenotype, but grows slowly in comparison to the wild type strain. Micro-array analysis (Barrass and Beggs, unpublished data) of a yeast strain with the proline-rich region deleted has a splicing defect, as it accumulated pre-mRNA in vivo when compared to the wild-type yeast strain. On the other hand, Prp8p N-terminal (aa2178-2310 in yeast) contains a Jab1/MPN domain that binds directly to ubiquitin (Bellare et al., 2006). Prp8p mutations in its Jab1/MPN domain (prp8-602 and prp8-603) not only reduced the binding of Prp8p to ubiquitin, but also reduced tri-snRNP formation (Bellare et al., 2006), suggesting a functional link between ubiquitin and the pre-mRNA splicing machinery.

Yeast Prp8p was found to cross-link to uniformly-labeled wild-type pre-mRNA substrates, but not to RNAs with mutations in the 5’ splice site or branch point that prevented spliceosome assembly (Whittaker and Beggs, 1991). To identify sites in pre-mRNA where Prp8p bound, cross-linking reagents were incorporated at specific sites in pre-mRNA substrates. Human Prp8p was found to contact the pre-mRNA (at -7, -3, -2, +3, but not -17 or -27, with respect to the 5’ splice site, review in Grainger and Beggs, 2005). Yeast Prp8p was also found cross-linked at 5’ splice site (at positions -1, -2, -8, but not at +4 in the intron, before step one, Teigelkamp et al., 1995). Photo-crosslinking analysis at the 3’ splice site showed that yeast Prp8p cross-linked to ‘U’ at the exonic 3’ splice site +1, but was not detected in a Prp2p-depleted extract that was stalled prior to step one. When purified yeast Prp2p was added back to the Prp2p-depleted reaction, the Prp8p-3’ splice site interaction was restored, indicating that this contact was initiated during or subsequent to step one (Teigelkamp et al., 1995).

In reconstituted yeast U5 and U6 snRNPs, the whole 5’ stem-loop system of yeast U5 snRNA, including the invariant loop 1 nucleotides U96-U99 and the internal loops IL1 and IL2 (Figure 1.6) were cross-linked to yeast Prp8p (Dix et al., 1998). Also, Prp8p cross-linked to U6 (at U54), which is in the conserved region that contributes to the catalytic core of the spliceosome (Vidal et al., 1999). Both U5 snRNA loop 1 and Prp8p can be cross-linked to the 5’ and 3’ splice sites in HeLa or yeast extracts, and these contacts occur at similar stages in the splicing reaction (Newman and Norman, 1992; Teigelkamp et al., 1995;
O'Keefe et al., 1996; Dix et al., 1998; Segault et al., 1999; Alvi et al., 2001; McConnell and Steitz, 2001; O'Keefe 2002). U5 snRNA was proposed to anchor and align the ends of the exons in spliceosome catalysis (O'Keefe et al., 1996). Surprisingly, U5 loop 1 is not essential for splicing in HeLa nuclear extract (Segault et al., 1999), and is only required for step two in yeast extracts (O'Keefe et al., 1996; O'Keefe and Newman 1998). Interestingly, cross-linking of yeast Prp8p to the 3' splice site occurred even in the absence of U5 loop1, (Dix et al., 1998) suggesting that Prp8p may anchors the exons in the spliceosome, acting in conjunction with U5 loop 1 to align the ends of the exons for the second trans-esterification reaction (Dix et al., 1998; O'Keefe and Newman 1998).

Newman and co-workers have developed a new proteolytic mapping approach to characterise contacts between yeast Prp8 and catalytic core RNAs (Turner et al., 2006). Short peptide tags, containing protease sites using a custom-built Tn5-based transposon, were randomly inserted into yeast Prp8p. This strategy allows regions of Prp8p that cross-linked to RNA to be narrowed down after TEV protease digestion. The mapping data from cross-linking analysis revealed the sites of RNA interactions in the central region of Prp8p (Figure 1.7).

**Figure 1.7. Regions of Prp8 that contact RNA in snRNPs and spliceosomes.**
Schematic representation of Prp8 protein and RNA crosslinks. Cross-linked residues are circled, short-range cross-links using uridine and 4-thio-uridine (green circle). Extensive Prp8p cross-links with U5 snRNA are also represented, as is the U6 residue 54. Both the U5 and U6 cross-links were only demonstrated in snRNPs. All numbering relates to *S. cerevisiae*. (Figure adapted from Grainger and Beggs, 2005, and regions of Prp8p that crosslink with 4-thio-uridine (orange and blue circles) were referred to Turner et al., 2006.).
1.10 Retinitis pigmentosa

Retinitis pigmentosa (RP) is a heterogeneous group of inherited bilateral retinal pigmentary dystrophies, which is linked primarily to night-blindness, followed by progressive loss of peripheral vision, reduced central visual sharpness and acquired colour vision defect (Peter Herse, 2005). The progressive demise of the photoreceptors also precipitates other pathological symptoms in the retina, including the attenuation of the retinal vasculature and the accumulation of intra-retinal pigment deposits, from which the disease gets its name (reviewed in Kennan et al., 2005).

So far, twelve loci have been linked to autosomal dominant forms (adRP) of the disease, and 4 of these genes are related to pre-mRNA splicing (Table 1.1). However, additional loci are yet to be identified.

Table 1.1 The known adRP genes

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Protein product</th>
<th>Locus name</th>
<th>Locus</th>
<th>Protein function</th>
<th>Expression profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA4</td>
<td>Carbonic anhydrase IV</td>
<td>RP17</td>
<td>17q22</td>
<td>Reversible hydration of carbon dioxide</td>
<td>Choroid and widespread</td>
</tr>
<tr>
<td>CRX</td>
<td>Cone-rod homeobox protein</td>
<td>19q13.3</td>
<td></td>
<td>Transcription factor</td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>FSCN2</td>
<td>Retinal fascin</td>
<td>17q25</td>
<td></td>
<td>Actin bundling in the connecting cilium</td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>IMPDH1</td>
<td>Inosine monophosphate dehydrogenase type 1</td>
<td>RP10</td>
<td>7q32.1</td>
<td>Guanine nucleotide synthesis</td>
<td>Photoreceptors &amp; widespread</td>
</tr>
<tr>
<td>NRL</td>
<td>Neural retina leucine zipper protein</td>
<td>RP27</td>
<td>14q11.2</td>
<td>Transcription factor</td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>PAP-1</td>
<td>Pim-1 associated protein 1</td>
<td>RP9</td>
<td>7p14.3</td>
<td>Pre-mRNA splicing</td>
<td>Photoreceptors &amp; widespread</td>
</tr>
<tr>
<td>PRPF3</td>
<td>Prp3p</td>
<td>RP18</td>
<td>1q21.1</td>
<td></td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>PRPF8</td>
<td>Prp8p</td>
<td>RP13</td>
<td>17p13.3</td>
<td></td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>PRPF31</td>
<td>Prp31p</td>
<td>RP11</td>
<td>19q13.4</td>
<td></td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>RDS</td>
<td>Retinal degeneration slow protein</td>
<td>RP7</td>
<td>6p21.1cen</td>
<td>Outer segments disc structure</td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>RHO</td>
<td>Rhodopsin</td>
<td>RP4</td>
<td>3q22.1</td>
<td>Photo-sensitive pigment</td>
<td>Photoreceptors</td>
</tr>
<tr>
<td>RP1</td>
<td>Retinitis pigmentosa RP1 protein</td>
<td>RP1</td>
<td>8q11-q13</td>
<td>Structure of Outer segments disc &amp; axonema</td>
<td>Photoreceptors</td>
</tr>
</tbody>
</table>

(Table adapted from Kennan et al., 2005)

The four known splicing factors genes that have been identified as being mutated in adRP: PRPF3, PRPF8, PRPF31, and PAP-1 (Pim-1 associated protein)) (Chakarova et al., 2002; McKie et al., 2001; Vithana et al., 2001; Martinez-Gimeno et al., 2003; Maita et al., 2004; Maita et al., 2005). All these factors were shown to be part of the tri-snRNP complex, and only PAP-1 has no homologue in budding yeast.
PRPF8 mutations causing adRP were found clustered in a 14-codon stretch within the last exon of the 7 kb transcript. The altered amino acid residues at the C-terminal exhibit a high degree of conservation across species as diverse as humans, Arabidopsis, trypanosome, and Saccharomyces cerevisiae (McKie et al., 2002).

Truncated PRPF31 proteins significantly inhibited pre-mRNA splicing of intron 3 of the RHO gene (Yuan et al., 2005). Analysis in cell culture has offered some evidence of reduced protein solubility for PRPF31 proteins that incorporate missense mutations (A194E and A216P). Both point mutations affecting PRPF31 proteins impede, but do not entirely prevent its translocation from cytoplasm to nucleus. A high proportion of the expressed mutant protein was insoluble, which could be due to the protein mis-folding, resulting in accumulation of insoluble aggregates in the cytosol (Deery et al., 2002).

The identification of PRPF8 and PRPF31 led Chakarova and co-workers to screen HPRP3 as a candidate for causing adRP. Indeed, two different missense mutations are clustered within a two-codon stretch in the 11th exon of the HPRP3 gene. One of the mutations (T494M) is seen repeatedly in apparently unlinked families, raising the possibility of a mutation hot spot (Chakarova et al., 2002). Adjacent to these mutations is the C-terminal region that is required for association of Prp3p with PAP-i. Both Prp3p and PAP-1 were found to be components of the U4/U6.U5 tri-snRNP complex (Maita et al., 2004). The H137L mutation of PAP-1 had no effect on splicing, while the other (D170G) resulted in both a defect in splicing activity and a decreased proportion of phosphorylated PAP-1 (Maita et al., 2004).

The mutations of splicing factors that relate to adRP suggest a link between spliceosome formation and adRP. Rather than these mutations causing a dominant-negative effect, the disease phenotype might relate to thresholds in the level of splicing. In other words, photoreceptors might survive until the level of splicing of rhodopsin, or other mRNAs, drops below a crucial level, at which point the cell would be irrevocably damaged, eventually leading to apoptosis.

1.11 U5 snRNP

The U5 snRNP is an active central component of both the U2- and U12-type spliceosomes (Grabowski and Sharp, 1986; Frilander and Steitz, 2001), and has the same protein composition in both spliceosomes (Will et al., 1999) (Table 1.2). The gene encoding the S. cerevisiae U5 snRNA, SNR7, was initially identified on the basis of a 9-nucleotide (nt)
sequence element (nt 93 to 101) (U5 loop 1) that is invariant in all U5 snRNA sequences (Patterson and Guthrie, 1987). Studies in vivo in HeLa (Segault et al., 1999) and in vitro in yeast argue that the U5 loop 1 sequence affects splice site selection, particularly for introns with non-ideal 5′ splice sites (Newman and Norman, 1991).

Figure 1.8. Yeast U5 snRNA secondary structure. (Figure adapted from Chanfreau et al., 1996).

U5 is proposed to contact the 5′ splice site prior to the first splicing step and then maintain its hold, later aligning the ends of the exons for ligation in the second step (Umen and Guthrie, 1995; O’Keefe and Newman 1998). The highly conserved U5 loop 1 cross-links to the upstream exon prior to the first step, and to both exons during the second step of splicing (O’Keefe and Newman, 1998). In the yeast in vitro system, some loop mutations prohibited the second step of splicing (O’Keefe et al., 1996; O’Keefe and Newman, 1998). However, the same region of U5 surprisingly proved to be dispensable for splicing in HeLa nuclear extract (Ségault et al., 1999). The 2 forms of U5 snRNA (S & L) (Figure 1.8) are generated by 3′-end processing by RNase III enzymes. The U5L is generated by Rnt1p, but the factor that is responsible for USS generation is so far unknown (Chanfreau et al., 1996). The reason for two forms of U5 snRNAs being recruited into the spliceosome still remains a
mystery, but yeast cells possessing only U5S were shown to be able to support cell growth (Seraphin et al., 1991).

Prp8p or h220K (in human) (Grainger and Beggs, 2005) contacts both the 5' and 3' splice sites in the pre-mRNA (Whittaker and Beggs, 1991; Teigelkamp et al., 1995). At least two forms of U5 snRNP have been isolated so far; the simple form of U5 snRNP that was isolated by two-step affinity purification. First, all snRNP is isolated by utilizing antibodies specific for the tri-methylguanosine (m3G) cap, then Aar2p with a deca-histidine tag at its C-terminus was purified with Nickel-NTA column (Gottschalk et al., 2001). A more complex form of U5 snRNP that was also purified by affinity chromatography using a strain that possesses a Brr2 protein tagged with a polyoma epitope at its carboxy-terminus (Stevens et al., 2001). The Aar2p-U5 snRNP consists of Aar2p, Prp8p, Snu114p, Sm proteins, and U5 snRNA. On the other hand, Brr2p-U5 snRNP consists of all the components of Aar2p-U5 snRNP (except Aar2p) (Table 1.2), in addition to Snu40p, Prp28p, and Dib1p. Of the six U5 snRNP-specific proteins (from the Brr2p-U5 snRNP), three are NTPases (Brr2p, Prp28p, and Snu114p).

Table 1.2 U5 snRNP-associated proteins

<table>
<thead>
<tr>
<th>Human homologue</th>
<th>Accession number</th>
<th>Sequence motifs</th>
<th>Yeast homologue</th>
</tr>
</thead>
<tbody>
<tr>
<td>U5-220K</td>
<td>gi3661610</td>
<td>Jab1/MPN</td>
<td>Prp8p</td>
</tr>
<tr>
<td>U5-200K</td>
<td>gi45861372</td>
<td>2 DEXH</td>
<td>Brr2p</td>
</tr>
<tr>
<td>U5-116K</td>
<td>gi41152056</td>
<td>G-domain</td>
<td>Snu114p</td>
</tr>
<tr>
<td>U5-102K</td>
<td>gi40807485</td>
<td>TRP repeats</td>
<td>Prp6p*</td>
</tr>
<tr>
<td>U5-100K</td>
<td>gi41327771</td>
<td>DEAD, RS</td>
<td>Prp28p</td>
</tr>
<tr>
<td>U5-52K</td>
<td>gi5174409</td>
<td>GYF domain</td>
<td>Lin1p</td>
</tr>
<tr>
<td>U5-40K</td>
<td>gi4758560</td>
<td>WD40</td>
<td>No homologue</td>
</tr>
<tr>
<td>U5-15K</td>
<td>gi5729802</td>
<td>Thioredoxin fold</td>
<td>Dib1p</td>
</tr>
</tbody>
</table>

*Prp6p is associated with S. cerevisiae tri-snRNP complex, but not with U5. (Table adapted from Will and Luhrmann, 2005).

Brr2p and Prp28p are members of the DEXD/H box family. The NTPases of the U5 snRNP are involved in the critical switch in which U1 is replaced by U6 at the 5' splice site. This is an important stage in spliceosome activation, and contributes to the fidelity of 5' splice site recognition. The two unwinding events, that disrupt U1:5' splice site base-pairing, and U4:U6 base-pairing, allow the 5' splice site, U6 and U2 catalytic core structure to form (Brow 2002). Under normal conditions, Prp28p is an essential protein in yeast but becomes dispensable when U1C protein (a factor that stabilises the U1:5' splice site interaction) is mutated. This implicates Prp28p in disrupting the U1C interaction, either directly acting
against the protein or by disrupting the U1:5' splice site helix that forms its site of interaction (Chen et al., 2001). Disruption of the U1:5' splice site interaction is accompanied by the release of U6 from U4. The U4/U6.U5 tri-snRNP contains the base-paired U4:U6 structure. In humans, h200K (Brr2p in yeast) is capable of unwinding RNA helices, including a base-paired U4:U6 complex (Laggerbauer et al., 1998). Furthermore, it has been shown that a mutation in the ATPase domain of Brr2p (brr2-1) inhibits the ATP-dependent disruption of U4/U6.U5 tri-snRNPs in yeast cell extracts (Raghunathan and Guthrie, 1998).

Snu114p is the sole GTPase identified in the spliceosome to date and is related to translation elongation factor EF-2 (Fabrizio et al., 1997). As a homolog of elongation factor G, Snu114p contains three domains (III-V) predicted to undergo a large rearrangement following GTP hydrolysis (Brenner and Guthrie, 2005). Studies using a Snu114p mutant that switched specificity from GTP to XTP showed that stalled spliceosome would only unwind the U4:U6 helices when supplied with hydrolysable XTP (Bartles et al., 2003), implying that Snu114p has a role either in unwinding of U4:U6 or, more probably, in controlling the action of Brr2p.

The human Prp8p (h220K) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase (h200K), h116K, and a WD-40 protein (h40K in humans, no homologue in yeast). The h116K and h40K show the strongest association to h220K, which is stable in 0.4 M sodium thiocyanate (NaSCN). On the other hand, the interaction between h220K and h200K was preserved under 0.2 M NaSCN centrifugation conditions (Achsei et al., 1998).

1.12 Links between pre-mRNA splicing and other mRNA processing

Messenger RNA production requires synthesis of a pre-mRNA by RNA Pol II and processing of a nascent pre-mRNA by 5' capping, splicing of introns, 3' cleavage/polyadenylation to make mature mRNA. All these reactions have been found to be coupled to one another, such that protein(s) involved in one step also participated in another (Bres et al., 2005; Gornemann et al., 2005; Proudfoot et al., 2002; Robert et al., 2002; Gunderson et al., 1998), suggesting a cross-talk among mRNA processing complexes.

Transcription and splicing. Although transcription and splicing were discovered independently, recent evidence has shown that these two reactions closely influence one another (Proudfoot et al., 2002). The finding that spliceosome U snRNPs interact with human transcription elongation factor TAT-SF1 (Cus2p in yeast; a splicing factor that is
associated with the U2 snRNP, Yan et al., 1998), and that the inclusion of splicing signals in nascent transcripts further stimulates transcription, suggests that, the recruitment of U snRNPs near elongating polymerase is important for transcription (Fong and Zhou, 2001). In addition, human splicing factor, SKIP (Prp45p in yeast), was also shown to associate with P-TEFb and enhances transcription elongation by HIV-1 Tat (Bres et al., 2005), supporting the fact that splicing factor(s) could serve in both transcription and splicing pathways. However, the feed back mechanism of efficient splicing to pre-mRNA transcription levels has so far not been demonstrated.

**Capping and splicing.** The first pre-mRNA processing step occurs after about 20-30 nucleotides have been synthesised. First, an RNA 5' triphosphatase (RTP) hydrolyses the triphosphate of the first nucleotide to disphosphate. Then, a guanylytransferase catalyses the fusion of a GMP moiety from GTP to the first nucleotide of the pre-mRNA via an unusual 5'-5' triphosphate linkage. Finally, a methyltransferase methylates the N7 position of the transferred GMP. This initial cap structure is recognised by the cap binding complex (CBC20p and CBC80p) (Proudfoot et al., 2002). Surprisingly both CBC20 and CBC80 are dispensable in yeast (Fortes et al., 1999). However, chromatin immunoprecipitation (CHIP) analysis showed that, in the absence of CBC20p and CBC80p, U5 snRNP binding to pre-mRNA was nearly abolished. This indicates that CBC is required for proper coupling of splicing to transcription in yeast (Gornemann et al., 2005).

**Polyadenylation and splicing.** The 3' end of almost all eukaryotic mRNAs comprises a homopolymer of 20-250 adenosine residues called the poly(A) tail. The poly(A) tail is added to pre-mRNA in the nucleus by cleavage and polyadenylation. The coupling of polyadenylation and splicing is integrally linked to exon definition (Berget 1995), particularly the definition of the last exon of a pre-mRNA, which involves protein-protein interactions across the exon, between splicing factors at the 3' splice site, and components of the polyadenylation complex at the polyadenylation signal (Cooke et al., 1999; Cooke and Alwine, 2002). In humans, U1 snRNP binding to the 5' splice site results in inhibition of polyadenylation via a direct interaction between U1 70K and poly(A) polymerase (Gunderson et al., 1998).
1.13 The Aim of the current studies

Prp8p is one of the most conserved proteins among eukaryotes, both in its sequence and in its large size (280 kDa). But Prp8p is unique, having no obvious homology to other proteins. Research so far has placed Prp8p as the only protein close to the catalytic centre of the splicing machinery. As it is unknown whether pre-mRNA splicing is catalysed by RNA(s) or protein(s), it is important to further investigate the role of this protein. Beside Prp8p, the proteins that are associated with Prp8p in the U5 snRNP (such as Snu114p and Brr2p) are also highly conserved. In addition, the fact that mutation of human Prp8p results in retinitis pigmentosa type 13 has emphasized the importance of further analysis of this protein.

In order to understand more about this protein, this study has employed protein purification methods combined with genetic dissection of Prp8p, to analyse the association of Prp8p with other splicing factors. This approach allowed me to confirm the association of Prp8p (aa436-770) with Snu114p that is independent of Aar2p, and that dissection of Prp8p at aa2173 disrupts the association between Prp8p and Aar2p.

The strategy was extended with a bioinformatics approach to identify possible domains of Prp8p. This led me to identify the functional nuclear localisation signal (NLS) of yeast Prp8p that is located at its N-terminus (aa96-117). Further analysis of the Prp8p NLS showed that, upon its deletion, Snu114p and the U5 snRNA were delocalised from the nucleus to the cytoplasm, but remained associated with Prp8p. Additionally, Aar2p showed increased association with deleted-NLS-Prp8p when compared to full-length Prp8p. However, for the other U5 snRNP component, Brr2p, nuclear localisation was not affected. This allowed me to propose that Prp8p associated with Snu114p, Aar2p, and U5 snRNA in the cytoplasm before entering the nucleus; and that Prp8p forms a complex with Brr2p in the nucleus.

By site-directed mutagenesis, I have created the yeast Prp8p mutants (rp) that resemble the human Prp8p mutations that cause retinitis pigmentosa type 13. Despite the fact that Prp8p-rp nuclear localisation was not affected by the mutations, Prp8p-rp and Brr2p association was disrupted. At the same time, the Prp8p-Aar2p association was increased. This finding is consistent with the previous results of deleted-NLS-Prp8p analysis. Despite the fact that the Prp8p-Snu114p-Aar2p complex that forms in the cytoplasm was able to enter the nucleus, this complex was increased due to the disruption of association between Prp8p-rp5 and Brr2p. In addition, the identification of importin Kap95p (Importin-β in human) that functions in U2, U5, and U6 snRNPs nuclear import supports the proposal that the U5 snRNP has a cytoplasmic phase.
Spp382p was co-selected with Prp8p or vice versa (Chapter 3 & 5; Boon et al., 2006). The characterisation of Spp382p showed that this protein is spliceosome-associated. The protein binds to U2, U5, U6 snRNAs, and excised intron, and interacts with Prp43p in the absence of other yeast proteins. Genetic-depletion of Spp382p results in excised intron accumulation, suggesting that this protein is needed for the late-stage of splicing. In Spp382p-depleted extract the U5 snRNP protein, Prp8p, is associated with 8-fold more U2 snRNA and glycerol gradient analysis of snRNPs also showed reduced levels of free U5 and U6 snRNPs, but accumulation of U4/U6 di-snRNP, suggesting a defect in recycling snRNPs when compared to non-depleted cell extract. Overall results suggest that Spp382p could be needed for disassembly of post-splicing complexes and for promotion of re-generation of a new round of splicing. However, more experiments are needed to prove this assumption.
Chapter 2
Materials and Methods

2.1 General reagents

Chemicals were purchased from the following sources, except where stated otherwise: Amersham, Fisher, Invitrogen, Roche, Sigma.

Restriction enzymes, DNA and RNA polymerases and other enzymes were purchased from Roche, or New England Biolab.

2.1.1 Bacterial and yeast growth media

All growth media were autoclaved and stored at room temperature or 4°C. For solid media, 2% (w/v) Bacto agar was added prior to autoclaving.

<table>
<thead>
<tr>
<th>Table 2.1 Bacterial and yeast growth media</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medium</td>
</tr>
<tr>
<td>Drop-out powder</td>
</tr>
<tr>
<td>Luria-Broth (LB)</td>
</tr>
<tr>
<td>Sporulation medium</td>
</tr>
<tr>
<td>YMGR Drop-out</td>
</tr>
<tr>
<td>YMM Drop-out</td>
</tr>
<tr>
<td>YPDA</td>
</tr>
<tr>
<td>YPGR</td>
</tr>
<tr>
<td>YPR</td>
</tr>
<tr>
<td>5-FOA</td>
</tr>
</tbody>
</table>
### 2.1.2 Commonly used buffers

#### Table 2.2 Commonly used buffers

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Solution / Litre</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGK buffer (0.5L)</td>
<td>10 ml 0.5M HEPES (pH7.0), 0.75 ml 1M MgCl₂, 50 ml 2M KCl, 50 ml Glycerol, 389.25 ml H₂O.</td>
</tr>
<tr>
<td>50x Denhardt's reagent (0.5L)</td>
<td>5 g of Ficoll (Type 400), 5 g of polyvinylpyrroline, 5 g of bovine serum albumin (Fraction V) dissolved in 500 ml H₂O. Solution was filtered and stored at -20°C.</td>
</tr>
<tr>
<td>Buffer B</td>
<td>0.1 M potassium phosphate pH7.4, 1.2 M sorbitol.</td>
</tr>
<tr>
<td>Dialysis buffer (1.5L)</td>
<td>60 ml 0.5M HEPES (pH7.0), 37.5 ml 2M KCl, 0.6 ml 0.5M EDTA, 300 ml Glycerol, 0.75 ml 1M DTT, 1.1 L H₂O, 1 protease inhibitor cocktail tablet (Sigma Cat no: 1873580).</td>
</tr>
<tr>
<td>GG buffer</td>
<td>20 mM HEPES (pH 7.0), 100 mM KCl, 0.2 mM EDTA</td>
</tr>
<tr>
<td>1x IPP₃₀₀</td>
<td>6 mM HEPES, pH 7.9, 0.15 M NaCl, 2.5 mM MgCl₂, 0.05 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>1x IPP₅₀₀</td>
<td>6 mM HEPES, pH 7.9, 0.3 M NaCl, 2.5 mM MgCl₂, 0.05 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>2x IPP₅₀₀</td>
<td>6 mM HEPES, pH 7.9, 0.5 M NaCl, 2.5 mM MgCl₂, 0.05 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>2x IPP₆₀₀</td>
<td>12 mM HEPES, pH 7.9, 0.3 M NaCl, 5 mM MgCl₂, 0.1 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>2x IPP₁₀₀₀</td>
<td>12 mM HEPES, pH 7.9, 0.6 M NaCl, 5 mM MgCl₂, 0.1 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>20x MOPS</td>
<td>1 M MOPS (3-N-morpholine) propane sulfonic acid, 1 M Tris, 2% (w/v) SDS, 20.5 mM EDTA.</td>
</tr>
<tr>
<td>NTN₃₀₀</td>
<td>150 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.1 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>NTN₅₀₀</td>
<td>300 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.1 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>NTN₁₀₀₀</td>
<td>500 mM NaCl, 50 mM Tris-HCl, pH 7.5, 0.1 % (v/v) Nonidet P40.</td>
</tr>
<tr>
<td>1x PBS</td>
<td>137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.5 mM KH₂PO₄, pH adjusted to 7.4 with HCl.</td>
</tr>
<tr>
<td>1x PBST</td>
<td>137 mM NaCl, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 1.5 mM KH₂PO₄, pH adjusted to 7.4 with HCl, 0.1 % (v/v) Tween-20.</td>
</tr>
<tr>
<td>20x SSC</td>
<td>3 M NaCl, 0.3 M sodium citrate, pH 7.0.</td>
</tr>
<tr>
<td>20x SSPE</td>
<td>3.6 M NaCl, 200 mM Na₂HPO₄, 20 mM EDTA, pH adjusted to 7.4 with NaOH.</td>
</tr>
<tr>
<td>50x TAE</td>
<td>2 M Tris, 50 mM EDTA, 5.7% (v/v) Acetic acid.</td>
</tr>
<tr>
<td>10x TBE</td>
<td>0.9 M Tris-borate, 20 mM EDTA.</td>
</tr>
<tr>
<td>10x TBST</td>
<td>0.5 M Tris-HCl, 1.5 M NaCl, 1% (v/v) Tween-20, pH adjusted to 7.5 with HCl.</td>
</tr>
<tr>
<td>10x Western Transfer buffer</td>
<td>200 mM Tris Base, 1.5 M Glycine.</td>
</tr>
</tbody>
</table>
2.1.3 Oligonucleotides

Oligonucleotides that were used in this study are listed below. All oligonucleotides were purchased from MWG Biotech. The abbreviation of YCPCR stands for yeast colony PCR. The name of primers, such as F-Aar2-703, refer to forward primer for the AAR2 gene at a position 703 nucleotides after the start codon.

### Table 2.3 Oligonucleotides

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5' – 3')</th>
<th>Notes or Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FISH Probes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U1Fish1</td>
<td>A/CACCAATTGGAATTGGTGCTCAACTCTCTCCA GGCAGAAGAACAAAGGCCCCCCCCAAAAATCGTTAA</td>
<td>5' FcT modified</td>
</tr>
<tr>
<td>U1Fish2</td>
<td>AA/ACTCCTGCAAAAATGGAACAGCTAGA GAAAAGTATGCTCAAAAAAAGATGCTCTCCTACAAAG</td>
<td>5' FcT modified</td>
</tr>
<tr>
<td>U2Fish1</td>
<td>TT/GGACATAAACCGCTCGAAAGACAGGGGAAG GTATGAGACAA</td>
<td>5' Cy3 modified</td>
</tr>
<tr>
<td>U2Fish2</td>
<td>AA/GAAGAGCGAGAAGAAATCAAACATAGAGGC GCCCCTCCGCCTCA</td>
<td>5' Cy3 modified</td>
</tr>
<tr>
<td>U3A</td>
<td>AT/TCAGTGCTCTTIGTAAGATGTAAGAGGAGTAG CATTCCATATAGAAATGA</td>
<td>5' Cy3 modified, M. Spiller, Beggs lab</td>
</tr>
<tr>
<td>U5Fish1</td>
<td>GG/GTAAAGAGCAGAGACATGTGACCGAC</td>
<td>5' Cy3 modified</td>
</tr>
<tr>
<td>U6Fish1</td>
<td>TG/ATGATCTCTGTGTTGCAATAATGACCAAAT GTCCA</td>
<td>5' Cy3 modified</td>
</tr>
<tr>
<td><strong>Northern Probes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G8102</td>
<td>CACCGCTTCGCAGCAGT</td>
<td>U1 snRNA</td>
</tr>
<tr>
<td>G8103</td>
<td>CT/ACAAGTTAGCTAAAGCCAAAAAG</td>
<td>U2 snRNA</td>
</tr>
<tr>
<td>G6568</td>
<td>AGGTATCCAAAAATTC</td>
<td>U4 snRNA</td>
</tr>
<tr>
<td>U5-7wtsnr</td>
<td>AAGTCCAAAAAATAATGGCAAGC</td>
<td>U5 snRNA</td>
</tr>
<tr>
<td>U6_knockout</td>
<td>ATCCTGATTTGTTCATAATGACCAAA</td>
<td>U6 snRNA</td>
</tr>
<tr>
<td><strong>AAR2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Aar2-703</td>
<td>GG/TAAAGAACCCCGGTACTGTTGTGCCTTCTTAT TACAAA GGCATCCATGAAAGAGAAG</td>
<td>YCPCR</td>
</tr>
<tr>
<td>F-Aar2-TAP</td>
<td>GAGCACAACCCCAATTGTTGCGGCTCTTAT TACAAA GGCATCCATGAAAGAGAAG</td>
<td>TAP cassette amplification</td>
</tr>
<tr>
<td>R-Aar2-TAP</td>
<td>ATCAAGCTCATAATAGTGTAGATGGCCGCCA CGATGATCGTTATACGTACTCATAAGGG</td>
<td>TAP cassette amplification</td>
</tr>
<tr>
<td><strong>Brr2</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-Brr2-142</td>
<td>GT/GATATTCCGCATCCC</td>
<td>YCPCR</td>
</tr>
<tr>
<td>F-Brr2-6100</td>
<td>ACTGATCAGCACTTGCG</td>
<td>YCPCR</td>
</tr>
<tr>
<td>R-Brr2-6764</td>
<td>AACCCACAGACCCTAAAGAC</td>
<td>YCPCR</td>
</tr>
<tr>
<td>F2-Brr2</td>
<td>TAT/CTTGAGCGAGATAAAGAGTGTCTCTTGA AAATGATGAAA CGGATCCCCCCGGTATTTAATTA</td>
<td>PFA6a-HIS3-13Myc amplification</td>
</tr>
<tr>
<td>R1-Brr2</td>
<td>TTA/TATATTGAAATATTCATTGAGATTATCCAGGACT AACAATGATTAGTATGGCTCGTTTTAAC</td>
<td>PFA6a-HIS3-13Myc amplification</td>
</tr>
<tr>
<td>F-Brr2-TAP</td>
<td>TAT/CTTGAGCGAGATAAAGAGTGTCTCTTGA AAATGATGAAAATCCATGAAAGAGAAG</td>
<td>TAP cassette amplification</td>
</tr>
<tr>
<td>R-Brr2-TAP</td>
<td>TATATTGAAAAATCATTGAGATTATCCAGGACT ACAAATGATTATATGACACTCACTATAGGG</td>
<td>TAP cassette amplification</td>
</tr>
</tbody>
</table>

**LSM8**
Lsm8f3: CGAAATGAGCATGTAATATGGGAAAAAGTGT TAA
Lsm8r2: TTAATATTATTATTATTATTATTATTATTACT ATTTAGCAACTATAGAATTCGAGCTCGTTAAC
XbaI-Lsm8f: CGTCTAGAATGTCGACATGGAACATCT AATAAAA

PRP8
-380-60PRP8-F: GATGAAAACGTGAAGGCTA
F-PRP8-988: GGTATACGACGCCACATTAGAG
F-PRP8-3310: ATACCGGAACTTGAAGGTTAC
F-PRP8-3523: GAAAGAAGATGAGGGAGGAGGA
F-PRP8-4182: TGGTGGTGTTGGTAGTAGT
F-PRP8-5296: ATGTATCCCTCCCAAAAC
F-PRP8-5802: CTGCACACAGAAAAAAGCTCCTT
F-PRP8-6335: AAGAAGCTGCAGGAAAAG
F-PRP8-6522: TGATTTCTGGAGAGAGAAG
F-PRP8-6876: AATTCTCCACACCCCGGAAT
R-PRP8-451: AGGCACATTCACGAGTAGAT
R-PRP8-1956: ACCATCTGCTAATCGAAAGG
R-PRP8-4154: GTTGAAATCTTGGTGCC
R-PRP8-4876: GCACCAAAAGACCCGACAT
R-PRP8-5655: AGATGCCCACTTCGTGGA
R-PRP8-5878: CTTGCGCCTGATTGCTA
R-PRP8-6689: AATAGGACGATACAGCAAACAGTC
R-PRP8-7233: TACATCGATTTTGCTGCCCT
R-PRP8-(+174): TTCTATGTCGGGAAAGCGGTAG
R1-PRP8: ATATCTATGAAATACGATTCCAGTTATGGAATATATATCGAATTTCCGACTCGTTAAC
F2-PRP8: GCGGGGGACGAAAGATGAGGCGCAGAAAAT
F4-PRP8: GTGCGGAGTCAGATACGACGCACTATAGGG
R4-(D78Prp8) Deletion (aa1-78 of Prp8p): TAATTCTCTTTTCTGTCTATTTCTCTCAATCTAGACCCACCGGAACAGATCCGATT
R4-Prp8: GCTGTTCTTCTACACATCATCACCATCATCAGAATTTCTGAGACCCGCCATGTCGATT
F-PRP8-439-TAP: TCCCTGTCTTATACACATCATCTACCAACATCCAGAGCATCCGATT
R-PRP8-439-TAP: CTTATGAAATAACAGATTCCAGTTATGGAAGATATCAGATACGACGACTACTAGGG
Prp8-Tap-F: GCCGGGGACGAAAGATGAGGCGCAGAAAAT
Prp8-Tap-R: TATATATCTATGAAATACGATTCCAGTTATGGAAGATATCAGATACGACTACTAGGG
Prp8-770-F: TCCTGAGTTCAGGAAAAGCTCGGTATGGGACAT
<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>Insertion/Deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTGTCGTATAGATCTTGTTCATTATTCTCAAGAAGTACGACTCACTATAG</td>
<td>pJU204-Rp5</td>
</tr>
<tr>
<td>TCTCAT</td>
<td>pJU204-Rp6</td>
</tr>
<tr>
<td>TFAGCG</td>
<td>pJU204-Rp7</td>
</tr>
<tr>
<td>SMB1</td>
<td></td>
</tr>
<tr>
<td>F-SMB1-ProtA</td>
<td>CCTAAACAAAGGAAGTTCGACCGCCCAACAGGGTTTTAGGAGAACAAAC</td>
</tr>
<tr>
<td>R-SMB1-ProtA</td>
<td>AAAAATATGCTACAAGTCGACTCACTATAG</td>
</tr>
<tr>
<td>F-SMB1-398</td>
<td>CCGGAAAAGATCGCTAAAC</td>
</tr>
<tr>
<td>SMD1</td>
<td></td>
</tr>
<tr>
<td>F-SMD1-ProtA</td>
<td>CGCGGATGACCGCACTAAGTGAAAGAAGGACAGGAGAC</td>
</tr>
<tr>
<td>R-SMD1-ProtA</td>
<td>AGTAATGCTGACGATAGCAACCCGAACAGCAGCAGTACGACTCACTATAGG</td>
</tr>
<tr>
<td>F-SMD1-373</td>
<td>GACCCCAGCAAAAAGAGA</td>
</tr>
<tr>
<td>SME1</td>
<td></td>
</tr>
<tr>
<td>F-SME1-ProtA</td>
<td>ATACCTGTTGAAAGGCAGAATATATCACATTTGTAATAACGCAAGGACAGGAGAC</td>
</tr>
<tr>
<td>R-SME1-ProtA</td>
<td>GTCACTTGGATATTGTCTTTCTGTGTTCCTCCCTTTTCTTTCTTACAGAGGAG</td>
</tr>
<tr>
<td>F-SME1-105</td>
<td>CGCCATAAGAATCAGGAG</td>
</tr>
<tr>
<td>Snu14</td>
<td></td>
</tr>
<tr>
<td>F-Snu14-2863</td>
<td>AAACCGCAGCTATCAAC</td>
</tr>
<tr>
<td>F-Snu14-TAP</td>
<td>AGCGCTGAAATTATAGCTCAATTTAAAGGAAGAAAT</td>
</tr>
<tr>
<td>R-Snu14-TAP</td>
<td>AATATTTGGACATATTGGCTTAATTCCTTATGCGC</td>
</tr>
<tr>
<td>CAAGATTTCATACGACTTCATATAGGG</td>
<td>amplification</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td>F-YKR022C/769</td>
<td>GAGAAAAATCTACCGCC</td>
</tr>
<tr>
<td>F-YKR022C-TAP</td>
<td>TTGGAAAAATCTAAGCAGCTTAAATAAAGCTCATTGCCAGGAAAAAGAAGAGG</td>
</tr>
<tr>
<td>R-YKR022C-TAP</td>
<td>TTATTATAGTACGGGATTTAAAATAACATGATTAACGGCTACCCACTATAGGG</td>
</tr>
<tr>
<td>YLR424W/SPP382</td>
<td>YLR424W amplification for pGFP-N-FUS insertion</td>
</tr>
<tr>
<td>F-Xba1-aa2YLR424W</td>
<td>TGCTCTAGAGGGATTCGGACTCAACCAC</td>
</tr>
<tr>
<td>R-Xho1-aa709YLR424W</td>
<td>CGCTCGAGCTAGAGTGCAAGGGCCCATAT</td>
</tr>
<tr>
<td>F1-424</td>
<td>CACCGAGAGGGTTAGAACCTTAAGCCTTCA</td>
</tr>
<tr>
<td>R1-424</td>
<td>ATATATATATCAGCTACCCATCTCTCTTTGAGATCCCGGCAGTCTGTTGGAATTC</td>
</tr>
<tr>
<td>F4-424</td>
<td>CACCGAGAGGGTTAGAACCTTAAGCCTTCA</td>
</tr>
<tr>
<td>R3-424</td>
<td>TTTGAAAAAGAACTTCTTTCTGTGTTGAGAGCCCATGATCGCACTCGAGTCTGGATCGTGTGTTGGAGTCC</td>
</tr>
<tr>
<td>F-YLR424W-TAP</td>
<td>TCCAGTGGGACCCCTTAGGCAAATTTATTGAGCGCCTCAGCTCCTCCATGGGAAAGAAGAGG</td>
</tr>
<tr>
<td>R-YLR424W-TAP</td>
<td>ATATATATATCAGCTACCCATCTCTCTTTGAGATCCCGGCAGTCTGTTGGAATTC</td>
</tr>
<tr>
<td>F-299-424W</td>
<td>TAGAACTAATAAGTGGAG CATCGG</td>
</tr>
<tr>
<td>F-1805-424W</td>
<td>TGGAAACAAGTAGTGACGCAG</td>
</tr>
<tr>
<td>R-549-424W</td>
<td>AACAACGGCTTCGCTAAAC</td>
</tr>
<tr>
<td>R-2565-45YLR424W</td>
<td>TTTAGTGGAGCTAAAATCCA</td>
</tr>
<tr>
<td>F-URA3</td>
<td>AGAGAATCAGCGCTCCCCAT</td>
</tr>
</tbody>
</table>

35
2.1.4 Plasmid vectors and constructs

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>M529</td>
<td>Modified pRS314; contains PRP8 coding sequence with N-terminal Protein A and C-terminal FLAG tag and interrupted by transposon with selective marker (KanMX4) after codon 436.</td>
<td>A. Newman, Cambridge</td>
</tr>
<tr>
<td>p283</td>
<td>Modified pGEM1: Contains ACT1 template under the regulation of the T7 promoter.</td>
<td>O’Keefe et al., 1996</td>
</tr>
<tr>
<td>pAW28</td>
<td>CEN-URA3</td>
<td>A. Newman, Cambridge</td>
</tr>
<tr>
<td>pBS1365</td>
<td>CEN-ProteinA-URA3</td>
<td>Puig et al., 1998</td>
</tr>
<tr>
<td>pBS1479</td>
<td>CEN-TAP-TRP1</td>
<td>Rigaut et al., 1999</td>
</tr>
<tr>
<td>pBS1539</td>
<td>CEN-TAP-URA3</td>
<td>Rigaut et al., 1999</td>
</tr>
<tr>
<td>pFA6a-His3MX6-PGAL1-3HA</td>
<td>CEN-HIS3-pGAL1-3HA</td>
<td>Longtine et al., 1998</td>
</tr>
<tr>
<td>pFA6a-kanMX6-6-PGAL1-GST</td>
<td>CEN-HIS3-pGAL1-GST</td>
<td>Longtine et al., 1998</td>
</tr>
<tr>
<td>pFA6a-His3MX6-3HA</td>
<td>CEN-HIS3</td>
<td>Longtine et al., 1998</td>
</tr>
<tr>
<td>pFA6a-natMX6</td>
<td>CEN-Nat</td>
<td>Goldstein &amp; McCusker, 1999</td>
</tr>
<tr>
<td>pGEX-4T-2-424</td>
<td>GST-Spp382p expression plasmid</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP-N-FUS</td>
<td>CEN-URA3-Pmet2-GFP</td>
<td>Niedenthal et al., 1996</td>
</tr>
<tr>
<td>pGFP-N-424</td>
<td>CEN-URA3-Pmet2-GFP-YLR424W</td>
<td>This work</td>
</tr>
<tr>
<td>pGFP-N-NLS</td>
<td>CEN-URA3-Pmet2-GFP-prp8aa96-117</td>
<td>This work</td>
</tr>
<tr>
<td>pJU204</td>
<td>CEN-HIS3</td>
<td>Umen and Guthrie, 1995</td>
</tr>
<tr>
<td>pJU204ANLS</td>
<td>CEN-HIS3</td>
<td>This work</td>
</tr>
<tr>
<td>pJU204-rp5</td>
<td>CEN-HIS3</td>
<td>This work</td>
</tr>
<tr>
<td>pJU204-rp6</td>
<td>CEN-HIS3</td>
<td>This work</td>
</tr>
<tr>
<td>pJU204-rp7</td>
<td>CEN-HIS3</td>
<td>This work</td>
</tr>
<tr>
<td>pJU204-Y1542A</td>
<td>CEN-HIS3</td>
<td>This work</td>
</tr>
</tbody>
</table>

2.1.5 E. coli strains

DH5α was routinely used for plasmid cloning and propagation steps.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL21</td>
<td>F', ompT, hsdSB, (rB, mB), gal, dcm (DE3)</td>
<td>Stratagene</td>
</tr>
<tr>
<td>DH5α</td>
<td>lacZM15 (lacZYA-argF) recA1 endA1 hsdR17(rK, mK') phoA supE44 thi gyrA96 relA1</td>
<td>Invitrogen</td>
</tr>
</tbody>
</table>
### Table 2.6 Saccharomyces cerevisiae strains

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMA38 (2n)</td>
<td>MATα, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100</td>
<td>B. Dujon, Institute Pasteur.</td>
</tr>
<tr>
<td>BMA38a</td>
<td>MATα, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100</td>
<td>Albers et al., 2003</td>
</tr>
<tr>
<td>BY4741</td>
<td>MATα, his3Δ1, leu2Δ0, met15Δ0, Ura3Δ0</td>
<td>Mouaikel et al., 2002</td>
</tr>
<tr>
<td>ANMD5</td>
<td>nmd5::nat, otherwise as BY4741</td>
<td>Anita Hopper</td>
</tr>
<tr>
<td>ΔTGS1</td>
<td>MATα, his3Δ1, leu2Δ0, met15Δ0, Ura3Δ0, YPL157W::KanMX4</td>
<td>Mouaikel et al., 2002</td>
</tr>
<tr>
<td>cse1-1 or</td>
<td>cse1-1 Mata, ura3-52, ade2-101, his3-11, 15trp1</td>
<td>Molly Fitzgerald</td>
</tr>
<tr>
<td>1706</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RG8T</td>
<td>PRP8::TAP-TRP1, otherwise as BMA38a</td>
<td>R. Grainger, Beggs lab</td>
</tr>
<tr>
<td>RG8T/2173</td>
<td>PRP8(1-2173):Tn:PRP8(2173-2413):TAP-TRP1, otherwise as BMA38a</td>
<td>R. Grainger, Beggs lab</td>
</tr>
<tr>
<td>KL8TG</td>
<td>HIS3-PGAL1-HA3:PRP8::TAP-TRP1, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL8H</td>
<td>PRP8::3HA-HIS3, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL8TG/436</td>
<td>His3-PGAL1-HA3:PRP8(1-439):Tn:PRP8(437-2413):TAP-TRP1, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL8H/436</td>
<td>PRP8(1-439):Tn:PRP8(437-2413):3HA-HIS3, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL8H/436T</td>
<td>PRP8(1-439):TAP-URA3:Tn:PRP8(437-2413):3HA-HIS3, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL8T/770</td>
<td>PRP8(1-770):Tn:PRP8(771-2413):TAP-TRP1, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL1</td>
<td>Kan-PGAL1-GST:prp8aaal-78, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL1-Snu114</td>
<td>SNU114::TAP-URA3, otherwise as KL1</td>
<td>This work</td>
</tr>
<tr>
<td>KL1-Brr2</td>
<td>BRR2::TAP-URA3, otherwise as KL1</td>
<td>This work</td>
</tr>
<tr>
<td>KL1-Aar2</td>
<td>AAR2::TAP-URA3, otherwise as KL1</td>
<td>This work</td>
</tr>
<tr>
<td>KL1-Lsm8</td>
<td>LSM8::TAP-URA3, otherwise as KL1</td>
<td>This work</td>
</tr>
<tr>
<td>KL2T</td>
<td>YKR022C::TAP-URA3, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL3</td>
<td>TRP1::PGAL1-GST:prp8, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL3-Brr2</td>
<td>Brp2::13Myc-Hph, otherwise as KL3</td>
<td>This work</td>
</tr>
<tr>
<td>KL3-SmB1</td>
<td>SMBI::ProteinA-URA3, otherwise as KL3</td>
<td>This work</td>
</tr>
<tr>
<td>KL3-SmD1</td>
<td>SMD1::ProteinA-URA3, otherwise as KL3</td>
<td>This work</td>
</tr>
<tr>
<td>KL3-SmE1</td>
<td>SME1::ProteinA-URA3, otherwise as KL3</td>
<td>This work</td>
</tr>
<tr>
<td>KL4F</td>
<td>YLR424W::nat, [pGFP-N-424], otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL4G</td>
<td>Kan-PGAL1-HA:YLR424W, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>KL4G2T</td>
<td>YKR022C::TAP-URA3, otherwise as KL4G</td>
<td>This work</td>
</tr>
<tr>
<td>KL4T</td>
<td>YLR424W::TAP-TRP1, otherwise as BMA38a</td>
<td>This work</td>
</tr>
<tr>
<td>PSY870</td>
<td>MATα, rnal-1, ade2, ade3, leu2, ura3, MAT a carrying pPS714</td>
<td>Koepp et al., 1996</td>
</tr>
<tr>
<td>PSY871</td>
<td>MATα, rnal-1, rsll-1, ade2, ade3, leu2, ura3, carrying pPS714</td>
<td>Koepp et al., 1996</td>
</tr>
<tr>
<td>W303a</td>
<td>MATα, ade2-1 trp1-1 leu2-3,112 his3-11,15 ura3-1</td>
<td>R. Rothstein</td>
</tr>
</tbody>
</table>
2.1.7 Antisera

Table 2.7 Antisera

<table>
<thead>
<tr>
<th>Antibody Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti-Prp8p aa2-35, Rabbit IgG, primary</td>
<td>This work</td>
</tr>
<tr>
<td>Anti-Aar2p Rabbit IgG, primary</td>
<td>Gottschalk et al., 2001</td>
</tr>
<tr>
<td>Anti-GST (Z-5) Rabbit IgG, primary</td>
<td>Santa Cruz Biotechnology</td>
</tr>
<tr>
<td>Anti-HA (Y-11) Rabbit IgG, primary (Western blotting)</td>
<td>Santa Cruz Biotechnology</td>
</tr>
<tr>
<td>Anti-HA high affinity (3F10) Rat IgG, monoclonal antibody, primary (Immunofluorescent staining)</td>
<td>Roche Applied Science</td>
</tr>
<tr>
<td>Anti-Myc (A-14) Rabbit IgG, primary (Western blotting)</td>
<td>Santa Cruz Biotechnology</td>
</tr>
<tr>
<td>Anti-Myc (9E10) Mouse IgG, primary (Immunofluorescent staining)</td>
<td>Abcam Ltd.</td>
</tr>
<tr>
<td>Anti-Nop1p Mouse IgG, primary</td>
<td>J. Aris</td>
</tr>
<tr>
<td>Anti-Protein A Rabbit IgG, primary</td>
<td>Sigma, P-3775</td>
</tr>
<tr>
<td>Anti-Prp43p Rabbit IgG, primary</td>
<td>Lebaron et al., 2005</td>
</tr>
<tr>
<td>Anti-Snu114p Rabbit IgG, primary</td>
<td>Bartels et al., 2003</td>
</tr>
<tr>
<td>Anti-TMG (K121) Mouse IgG, primary</td>
<td>Calbiochem, NA02</td>
</tr>
<tr>
<td>Goat Anti-Mouse IgG Alexa Fluor 594 conjugated, secondary</td>
<td>Molecular Probes, A-11005</td>
</tr>
<tr>
<td>Goat Anti-Mouse IgG, Highly cross-absorbed Alexa Fluor 488 conjugated, secondary</td>
<td>Molecular Probes, A-11029</td>
</tr>
<tr>
<td>Goat Anti-Mouse IgG Peroxidase conjugated, secondary</td>
<td>Amersham Biosciences, NA931</td>
</tr>
<tr>
<td>Goat Anti-Rabbit, IgG Highly cross-absorbed Alexa Fluor 594 conjugated, secondary</td>
<td>Molecular Probes, A-11307</td>
</tr>
<tr>
<td>Goat Anti-Rabbit Peroxidase conjugated, secondary</td>
<td>Amersham Biosciences, NA934V</td>
</tr>
<tr>
<td>Goat Anti-Rat, IgG Alexa Fluor 594 conjugated, secondary</td>
<td>Molecular Probes, A-11007</td>
</tr>
<tr>
<td>Peroxidase-anti-peroxidase (PAP) Rabbit IgG, used in direct-immuno Western blotting</td>
<td>Sigma, P-2026</td>
</tr>
</tbody>
</table>

2.1.8 Markers

All size markers were obtained from Invitrogen life technologies:

<table>
<thead>
<tr>
<th>Marker</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 kb plus DNA Ladder</td>
<td>100 bp – 12 kb</td>
</tr>
<tr>
<td>BenchMark™ Protein Ladders</td>
<td>10 – 220 kDa</td>
</tr>
<tr>
<td>SeeBlue® Plus2 Pre-stained Standard</td>
<td>4 – 250 kDa</td>
</tr>
<tr>
<td>MagicMark™ Western Protein Standard</td>
<td>20 – 120 kDa</td>
</tr>
</tbody>
</table>
2.2 Microbiological methods

Propagation and storage of bacteria
Growth and storage of *E. coli*, preparation of competent *E. coli* and transformation were followed as in Sambrook and Russell (2001).

Propagation, storage, and transformation of yeast
Yeast was grown and stored as described in Methods in Yeast Genetics (2000). Yeast transformation was carried out by the Lithium Acetate method (Gietz *et al.*, 1991).

2.3 Nucleic acid methods

Many standard DNA and RNA manipulations were carried out either according to Sambrook and Russell (2001) or manufacturers’ instructions. These include: quantification of nucleic acids by spectrophotometric; deproteinisation of nucleic acids solution with phenol/chloroform/Iso-amyI alcohol (25:24:1) (P/C/I); precipitation of nucleic acids with ethanol; restriction endonuclease digestion of DNA; agarose gel electrophoresis of DNA; isolation of plasmid DNA from *E. coli*: isolation of DNA fragments from agarose gels; filling recessed 3’ ends of DNA; dephosphorylation of DNA; ligation of DNA fragments.

2.3.1 Yeast genomic DNA preparation

Yeast genomic DNA was prepared using a modified method based on Hoffman and Winston (1987). Ten ml yeast culture was grown in YPDA overnight at 30°C in a shaking incubator, then sedimented by centrifugation and resuspended in 10 ml SE buffer (0.9 M sorbitol, 0.1 M EDTA, pH8.0). Again, cells were sedimented by centrifugation and resuspended in 0.5 ml SE buffer. 50 μl of 20 mg/ml lyticase (Sigma) was added into the SE buffer and incubated 30 min at 30°C on a rotating wheel. The spheroplasted cells were spun down (3500 rpm, 5 min) and resuspended in 500 μl of Lysis buffer (0.1M Tris-Cl, pH8.0, 50 mM EDTA, 1% SDS). 200 μl glass beads (150-212 μm in diameter) and 32 μl of 4 M NaCl were added. Then the tube was vortexed for 30 s and subsequently spun down at 3500 rpm for 2 min. The liquid was purified twice by P/C/I extraction. The DNA was precipitated by adding 900 μl ice-cold 100% ethanol, and then pelleted by centrifugation, and subsequently washed with 70% ethanol. The pellet was air dried and resuspended in 50 μl TE buffer.
2.3.2 Extraction of RNA from yeast

RNA was isolated according to Schmitt et al. (1990). Ten ml culture was grown in YPDA to an OD_{600} of 0.25-0.5, the cells were harvested by centrifugation and resuspended in 400μl of AE (50mM Na acetate pH 5.3, 10mM EDTA) buffer. The resuspended cells were transferred to a 1.5 ml microcentrifuge tube and 40 μl of 10% SDS was added. The suspension was vortexed, and then an equal volume of fresh phenol was added. The mixture was again vortexed and incubated at 65°C for 4 min. After that, the mixture was rapidly chilled in a dry ice/ethanol bath until phenol crystals appeared, and then centrifuged for 2 min at maximum speed in a micro-centrifuge to separate the aqueous and phenol phases. The upper, aqueous phase was transferred to a fresh micro-centrifuge tube and extracted with phenol/chloroform Iso-amyl alcohol (25:24:1) at RT for 5 minutes. The extracted aqueous phase was then brought to 0.3 M Na acetate, pH 5.3, by adding 40 μl of 3 M Na acetate pH 5.3, after which 2.5 volumes of ethanol from total volume were added to precipitate the RNA. After washing with 80% ethanol, the pellet was dried and resuspended in 20 μl of sterile water and stored at −70°C until used.

2.3.3 Radioactive ladders preparation

The reaction mix was assembled in a 1.5 ml eppendorf tube in the following order: 1 μg DNA pBR322 Msp I digested (NEB, N3032S), 2 μl dNTP mix (A, T, G in 2.5 mM [ ]), 2 μl 20 μCi α[32P]dCTP, 1 μl 10X Klenow buffer, 1 μl DNA polymerase I large (Klenow) fragment (Promega, M2201), 3 μl H₂O and incubated at RT for 10 mins. The reaction mix was made up to 100 μl with H₂O. Subsequently, 33 μl Ammonium Acetate (10 M) and 0.5 μl tRNA were added and then subjected to Phenol/Chloroform/Isoamyl-alcohol precipitation. The precipitated nucleic acids were re-suspended in 50 μl TE buffer (10 mM Tris-HCl (pH 8.0), 1 mM EDTA).
2.3.4 *In vitro* transcription

DNA was digested by appropriate restriction endonucleases. After overnight digestion, followed by inactivation, complete DNA digestion was confirmed by agarose gel electrophoresis, and then the DNA was purified by PCR purification kit (Qiagen). The transcription mix was assembled in a 1.5 ml eppendorf tube with the following order:

1) Linearised template DNA 2 µl, 0.01 ug/µl
2) Distilled H₂O 6 µl
3) 5mM rNTP solution 2 µl
   - rNTP solution containing 5mM of rGTP, rCTP and rUTP and 0.2 mM rATP
4) 100 mM DTT (dithiothreitol) 2 µl
5) 10X transcription buffer 2 µl
   - 400 mM Tris-Cl (pH 7.5), 60 mM MgCl₂, 20 mM spermidine HCl and 50 mM NaCl
6) 1 µl 10.0 mM m⁷GpppG 1 µl
7) 2 µl 10 mCi/ml (α-³²p) rATP. 2 µl

The components were mixed by gently tapping the outside of the tube. Then the following were add:

8) Placental RNase inhibitor (10units) 1 µl
9) T7 RNA polymerase 1 µl

The reaction mixture was incubated for 30 minutes at 37°C, and terminated by putting the reaction mix on ice. The reaction mixture was diluted with distilled H₂O to 220 µl, then purified twice by P/C/I extraction and ethanol precipitation (in ammonium acetate with final concentration of 2.0-2.5 M). The transcript RNA was resuspended in 10 µl of RNase-free H₂O and 10 µl 100% formamide was then added. Full-length transcripts were purified by running on a 6% (w/v) denaturing acrylamide gel, the full length transcript gel band was cut out (visualised by autoradiography) and eluted using a unidirectional electrolutor (model UEA; IBI) according to the manufacturer's instructions.
2.3.5 *In vitro* splicing

A splicing reaction mixture was prepared as follows:

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) 0.6 M KPO₄, pH 7.0</td>
<td>1 µl</td>
</tr>
<tr>
<td>2) 25 mM MgCl₂</td>
<td>1 µl</td>
</tr>
<tr>
<td>3) 20 mM ATP</td>
<td>1 µl</td>
</tr>
<tr>
<td>4) 30% (w/v) PEG₈₀₀₀</td>
<td>1 µl</td>
</tr>
<tr>
<td>5) Splicing extract</td>
<td>4 µl</td>
</tr>
<tr>
<td>6) Radiolabelled transcript (=1000 cpm)</td>
<td>2 µl</td>
</tr>
</tbody>
</table>

The reaction mixture was incubated at 24°C for 25 min and terminated by placement on ice. Three µl of Proteinase K solution (1 mg/ml Proteinase K, 50mM EDTA, 1% (w/v) SDS) was added to the reaction mixture and the sample was incubated at 37°C for 45 min. Subsequently, 200 µl of splicing cocktail (50 mM NaOAc, pH 5.3, 1 mM EDTA, 0.1% (w/v) SDS, 25 µg/ml *E. coli* tRNA) was added and followed by 2 times P/C/I extraction and one ethanol precipitation (by placing in −70°C for 30 min), and centrifuged at 14000 rpm for 20 min at 4°C. Then the sample was washed with 70% ethanol.

The RNA pellet was dried at RT and resuspended in 5 µl distilled water, subsequently 5 µl 1 X loading buffer (100% formamide + bromophenol and cyanole blue) was added and heated at 90°C for 3 min before loading on 7% (w/v) denaturing polyacrylamide gel at a setting of 24 W for 1 h in 1 X TBE and visualised by autoradiography.
2.3.6 Co-immunoprecipitation of snRNAs/spliceosome

Preparation of antibody-bound PAS (Protein A Sepharose, CL-4B) beads preparation: 35 μl of pre-hydrated protein A Sepharose beads were washed 3 times in 1 ml NTN. The beads were mixed with 300 μl of NTN and incubated with appropriate amount of antibody on a rotating wheel at 4 °C overnight. Incubated solution was taken off, antibody-bound beads were washed 3 times in 1 ml of NTN and were then incubated in 100 μl blocking solution (100 μg/ml BSA, 100 μg/ml Glycogen, 100 μg/ml tRNA) for 1 hour on a rotating wheel at room temperature. Then the beads were washed 4 times with 1 ml NTN and once with 1x IPP150 (6 mM HEPES, pH 7.9, 150 mM NaCl, 2.5 mM MgCl₂, 0.05 % (v/v) Nonidet P40). The 1x IPP150 was taken off and 50 μl of 2 x IPP300 (12 mM HEPES, pH 7.9, 300 mM NaCl, 5 mM MgCl₂, 0.1 % (v/v) Nonidet P40) and 0.5 μl RNAsin were added to the beads.

This was followed by immunoprecipitation. Fifty μl splicing extract (for snRNA co-IP) or 50 μl splicing reaction mix (for spliceosome co-IP) was mixed with IgG-agarose or antibody-bound PAS (in 50 μl 2 x IPP300 and 0.5 μl RNAsin). Samples were incubated on a rotating wheel for 2 hours at 4 °C. After that, the beads were washed twice with 1 ml of NTN and once with 1 ml NT (150 mM NaCl, 50 mM Tris-Hcl, pH 7.5). The buffer was taken off, then 50 μl of proteinase K digestion mix (50 mM Tris-Hcl, pH 7.5, 5 mM EDTA, 1.4 % (w/v) SDS, 2 μg/μl Proteinase K) was added to the beads, followed by 30 minutes incubation in 37 °C. Beads with bound immunoprecipitates were subjected to P/C/I extraction and ethanol precipitation. Pellets were vacuum dried and resuspended in RNase free water, then samples were electrophoresed on a 6% polyacrylamide gel and analysed by Northern blotting.
2.3.7 Northern blot analysis of snRNAs

The RNA was transferred electrophoretically to Hybond-N nylon membrane at 60 V for 1 hour in 0.5 x TBE. Following transfer, RNA was cross-linked to membrane by Stratagene UV Stratalinker with “Autocrosslink” to immobilise the RNA on the membrane.

**End-labelling of oligonucleotides** – Oligonucleotides were labelled in a 20 μl reaction mixture by T4 polynucleotide Kinase (PNK) at 37 °C for 45 minutes. The reaction mixture was as follows: Oligonucleotide 10 pmol/μl (1μl), 10 x PNK buffer (2 μl), T4 PNK (10 U/μl) (0.5 μl), [γ-32P] ATP (3000 Ci/mmol) (2 μl), 14.5 μl distilled H2O.

End-labelled oligonucleotides were totaled up to 1 ml with hybridisation buffer (mix 15 ml of 20 x SSPE with 5 ml of 50 x Denhardt’s reagent, 2.5 ml of 10 % SDS and 1 ml of 10 mg/ml ss-DNA, and the mixture solution was made up to 50 ml with 24 ml of distilled H2O, giving a final concentration of 6x SSPE, 5x Denhardt’s reagent, 0.5 % SDS and 0.2 mg/ml ss-DNA) and was passed through 2 μm filter. The filter was rinsed through with 2 x 1 ml hybridisation buffer.

**Hybridisation of 5’ End-labelled Oligonucleotide Probes to RNA**; Purified end-labelled oligonucleotides (from 20 μl reaction mixture) were hybridised to RNA on the nylon membrane in 50 ml hybridisation buffer, at 37°C overnight. The membrane was washed 4 times with 6 x SSPE (pre-warm 6 x SSPE at 42°C for 20 minutes each). The results were visualised by autoradiography.

2.4 Protein Methods

2.4.1 Crude protein extract preparation for Western blotting

An overnight yeast culture was measured at OD600 and a total of 3 OD were spun down. The cell pellet was resuspended in 0.5 ml of 0.2 M NaOH and left on ice for 10 min. After that, 55 μl of 50% TCA was added to the mixture which was left on ice for 10 min. The pellet was spin down at 14000 rpm for 1 min. The cell pellet was resuspended in 35 μl of dissociation buffer (100 mM Tris buffer (pH 6.8), 4 mM EDTA (pH 8.0), 4% SDS, 20% glycerol, 2% β-mercaptoethanol and trace amount of bromophenol blue), followed by adding 15 μl of 1 M Tris buffer (non-adjusted pH). Then the sample was heated at 95°C for 10 min. The sample was ready for loading on a gel or stored at −20°C.
2.4.2 Yeast splicing extract preparation

200 ml of overnight yeast culture was used as a starter culture. Six litres of medium were inoculated with starter culture to OD_{600} approximately 0.1, and cells were harvested when cell density reached 0.8 OD_{600} by centrifugation in JLA 10.500 rotor (5000 rpm, 5 min, 4°C). The cell pellet was resuspended in 50 ml of cold AGK buffer and sedimented by using Mistral 1000 centrifuge at 3500 rpm, 5 min, RT. This step was repeated by resuspending the cells pellet in 30 ml cold AGK buffer.

The supernate was removed and yeast pellet was weighed. Cells were resuspended in 0.4 volume of the cell pellet weight (i.e. 4 ml for 10 g of cells) in cold AGK buffer, and DTT was added to a final concentration of 2 mM. The yeast suspension was then frozen by adding drop-wise into liquid nitrogen. The frozen cell pellets could be either processed immediately or stored for up to 12 months at —80°C. The cell pellets were then added to a mortar containing liquid nitrogen and ground for 20-25 mins with occasional additions of liquid nitrogen to ensure the pellet did not thaw. The frozen powder was transferred into a pre-chilled beaker and thawed on ice. The completely thawed cell powder was stirred gently on ice for 2-3 minutes, then the supernate was transferred into a pre-chilled polycarbonate centrifuge tube and centrifuged in a JA 25.50 rotor (17000 rpm, 30 min, 4°C).

The supernate was then transferred to a pre-chilled polycarbonate centrifuge tube. The top layer of the supernatant containing lipid was avoided as much as possible during the transfer. This was then centrifuged in a 70.1 Ti rotor (40000 rpm, 60 min, 4°C). The resulting supernate was transferred to dialysis tubing with a molecular weight cut-off of 10 kDa. This was placed in 1.5 L of pre-chilled dialysis buffer and dialysed for 3 hours at 4°C. It was then placed in pre-chilled Eppendorf tubes and snap frozen in liquid nitrogen. The frozen tubes were stored at —80°C until use.
2.4.3 Co-immunoprecipitation from yeast protein extracts

Antibody-bound PAS (Protein A Sepharose, CL-4B) beads preparation: 30 µl of pre-hydrated protein A Sepharose beads were pelleted by 4 seconds pause centrifugation and washed in 1 ml NTN₁₅₀ 3 times. The beads were mixed with 300 µl of NTN₁₅₀ and incubated with appropriate amount of antibody (15 µl for 8.6 antibody) on a rotating wheel at 4°C for 2 hours or overnight. Incubated solution was taken off and antibody bound beads were washed twice in 1 ml of NTN₁₅₀ and once in 1 ml 1x IPP₁₅₀.

For IgG agarose beads, the beads were washed twice in 1 ml NTN₁₅₀ and 1 ml 1x IPP₁₅₀, then these steps were as followed: 50 µl of 2x IPP₃₀₀ and 50 µl splicing extracts were added to the beads and incubated at 4°C for 2 hr on a rotating wheel. The beads were washed 3 times with 1 ml 1x IPP₁₅₀, and the final 1x IPP₁₅₀ solution was removed carefully. The pellet was resuspended in 20 µl 1 x SDS loading buffer. The sample was vortexed and heated at 95°C for 5 minutes. The samples were used for SDS-PAGE and Western blotting.

2.4.4 TAP tag purification

The TAP tag purification protocol is according to the Séraphin Lab’s method (Rigaut et al., 1999), with some modifications. The splicing extract preparation was according to section 2.4.2.

IgG column preparation: First, the bottom of the column (Econo-column, Bio-Rad) was snap-opened, then 200 µl of IgG agarose bead suspension (SIGMA A2909) was added to the column and washed with 5ml of IPP₁₅₀ (10mM Tris-Cl pH 8.0, 150mM NaCl, 0.1% NP40). The bottom of the column was subsequently closed with a stopper.

Binding of IgG beads: Before adding the splicing extract to the column, each ml of splicing extract was adjusted with 10 µl of 1M Tris-Cl pH 8.0, 25 µl 4M NaCl, and 10 µl 10% NP40 (to give a final concentration of 10 mM Tris-Cl pH 8.0, 100 mM NaCl, and 0.1% NP40). After that, 2 ml of splicing extract were added into the column and the top of the IgG column was closed. Subsequently, the IgG column was rotated for 2 hours at 4°C. After that, the top plug of the column was removed, followed by removal of the bottom plug to allow the IgG column to drain by gravity flow. The drained IgG column was washed with 60 ml IPP₁₅₀ and 20 ml TEV cleavage buffer (10 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.1% NP40, 0.5 mM EDTA, 1 mM DTT).
TEV cleavage: The bottom of the IgG column was closed and 1 ml of TEV cleavage buffer and 60 Unit of TEV enzyme (Invitrogen) were added. The top of the IgG column was closed and rotated for 2 hr at 18°C.

Calmodulin column preparation: While waiting for the 2 hr incubation time above, a new column bottom was snap-opened and then 200 μl calmodulin bead suspension was added (STRATAGENE, #214303). Then the column was washed with 5 ml CBB (10 mM Tris-Cl pH8.0, 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM CaCl2, 0.1% NP40, 10mM BME) and the bottom of the calmodulin column was closed with a plug. After 2 hr, the top and bottom plugs of the IgG column were removed and the eluate from the IgG column was collected into the calmodulin column by gravity flow. The solution remaining in the IgG column dead-volume may be eluted with an additional 200 μl of TEV cleavage buffer, this gives a total of 1.2 ml elution.

Binding to, and elution from, Calmodulin beads: To the previous 1.2 ml eluate, 3 volumes of CBB (3.6ml) and 4.8 μl of 1M CaCl2 were added into the calmodulin column to give a final concentration of 1 mM CaCl2. The top of the calmodulin column was closed, and rotated for 2 hour at 4°C. After binding, the liquid was allowed to drain by gravity flow. The calmodulin column was then washed with 60ml CBB. After that, the bottom of the calmodulin column was closed and the complex bound to the beads was incubated 10 mins with CEB (Calmodulin elution buffer; 10 mM Tris-Cl pH8.0, 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 50 mM EGTA, 0.1% NP40, 10 mM BME). The elute was then collected in a 1.5 ml eppendorf tube.

TCA precipitation: The eluate was adjusted to 20% TCA with addition of 500 μl of 60% TCA, mixed and placed on ice for 30 min. The precipitate was spun down at 14000g for 30 min at 4°C. The supernate was removed and the pellet was washed once with 200 μl ice-cold acetone and twice in 200 μl ice-cold 80% ethanol (14000g for 5 min at 4°C). Wash solutions were removed and the pellet was air-dried. The pellet was resuspended in 20 μl 1x SDS loading buffer (sample can be stored at -20°C at this point).

2.4.5 Production and purification of GST-tagged recombinant protein from E. coli

LB broth (5 ml) containing 100 μg/ml ampicillin was inoculated with a single colony of E. coli (BL21) containing the expression plasmid and grown overnight at 37°C with agitation. This was then used to inoculate 40 ml of LB broth with 100 μg/ml ampicillin. This culture was then grown at 37°C to an OD600 of 1.6-1.7. To induce expression of the recombinant
protein, IPTG was added to a final concentration of 0.75 mM. The culture was then incubated for 2 days at 15°C. Cells were then harvested (3500 rpm, 5 mins, 4°C).

**Preparation of protein extract:** Cells were resuspended in 4 ml of Lysis Buffer A (50 mM Tris-HCl (pH7.5), 250 mM NaCl, 10% (w/v) Sucrose). Lysozyme was added to a final concentration of 0.2 mg/ml and the extract was stirred on ice for 45 mins. Triton-X-100 was added to a final concentration of 0.1% (v/v) and the extract stirred on ice for a further 5 mins before being transferred to a polypropylene centrifuge tubes and centrifuged (10000 rpm, 10 mins, 4°C; Beckman JA25.5 rotor). The supernatant was transferred into an eppendorf tubes and stored at -80°C until required.

**Purification of GST-tagged Recombinant Protein from E. coli:** 200 μl of Glutathione Sepharose 4B (Amersham) were washed 3 times with 10 ml of 1X PBS in poly-prep chromatography column (Biorad, catalog 731-1550). The supernate was then removed and 2 ml of the thawed protein extract was added before incubation (1 h, 4°C). The supernate was drained off from the column. Then, the beads were washed 3 times with 10 ml 1X PBS before incubated with 200 μl Elution buffer (20 mM Glutathione in 50 mM Tris-HCl pH8.0) for 10 mins at RT. The eluate was placed in dialysis unit (Slide-A-Lyzer, 7 KDa cut-off, Pierce) and dialysed against 500 ml of BC100 (20mM Tris-HCl, pH8.0, 100 Mm KCl, 20% (w/v) glycerol) for 3 h. After dialysis, protein extract was kept in an eppendorf tube and stored at -80°C).

2.4.6 Co-immunoprecipitation of E. coli recombinant protein

Recombinant protein (4 μg of GST-Spp382p) purified from E. coli (section 2.4.5) was mixed with an equal amount of 10His-Prp43p (4 μg) and incubated in 1X IPP150 buffer (section 2.1.2) in the total volume of 100 μl. The protein mix was then incubated with 20 μl pre-washed (3 x 1 ml NTN150 buffer and 1 x IPP150 1 ml) glutathione-sepharose beads for 2 hr on a rotating wheel at 4°C. The protein mix (supernatant) was then removed and kept. The glutathione-sepharose beads were washed two times with 1 ml NTN150 buffer and 1 time 1 ml NT150 buffer. The glutathione-sepharose beads were then resuspended in 25 μl of 1 x SDS buffer. Samples were then incubated in 65°C for 10 mins, fractionated in 4-12% PAGE gel and Western blotted with desired antibodies.
2.5 Primer design, PCR and DNA sequencing

Yeast colony PCR Primers were designed by using primer design software in *S. cerevisiae* Database (http://seq.yeastgenome.org/cgi-bin/web-primer). The optimum primer melting temperature was selected as 50°C. The optimum GC content selected was 50%.

The 20 µl PCR reaction was as follows: 2 µl 10X LA PCR buffer II (Magnesium\(^{2+}\) plus) from TAKARA, 3.2 µl dNTPs mix (dATP, dCTP, dGTP and dTTP, 2.5 mM each), 0.20 µl Forward primer (100 pmol/µl), 0.20 µl Reverse primer (100 pmol/µl), 14.2 µl H\(_2\)O, trace amount of yeast colony, 0.2 µl Roche Taq. The PCR mixture was run in MJ Research, PTC-200 PCR machine using the PCR program as follows:

<table>
<thead>
<tr>
<th>Step</th>
<th>Reaction</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Denaturing</td>
<td>95°C</td>
<td>6 mins</td>
</tr>
<tr>
<td>2</td>
<td>Denaturing</td>
<td>94°C</td>
<td>20 seconds</td>
</tr>
<tr>
<td>3</td>
<td>Annealing</td>
<td>50°C</td>
<td>50 seconds</td>
</tr>
<tr>
<td>4</td>
<td>Extension</td>
<td>72°C</td>
<td>1 min/kb</td>
</tr>
<tr>
<td>5</td>
<td>Final extension</td>
<td>72°C</td>
<td>6 mins</td>
</tr>
</tbody>
</table>

34 cycles of Step 2 to Step 4

DNA sequencing

Sequencing reactions were performed with the BigDye Terminator Cycle Sequencing Ready Reaction. The sequencing mix was set up as follows:

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Template DNA</td>
<td>2 µl</td>
</tr>
<tr>
<td>Primer</td>
<td>1 µl (1.6 pmol/µl)</td>
</tr>
<tr>
<td>BIG Dye</td>
<td>4 µl</td>
</tr>
<tr>
<td>Water</td>
<td>3 µl</td>
</tr>
</tbody>
</table>

The following sequencing program was run for 40 cycles:

<table>
<thead>
<tr>
<th>Step</th>
<th>Reaction</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96°C for 30 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>50°C for 15 seconds</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>60°C for 4 mins</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sequencing samples were run by the SBS Sequencing Service, Ashworth Laboratories, University of Edinburgh.
2.6 Site-directed mutagenesis by whole plasmid synthesis

The oligonucleotide primers, each complementary to the opposite strand of the vector and containing the desired mutation(s) or deletion, are extended during temperature cycling by PFU DNA polymerase (Figure 2.1). The 50 μl PCR reaction was as follows: 5 μl 10x PFU (High fidelity) buffer, 2.5 μl dNTPs mix (dATP, dCTP, dGTP and dTTP, 10 mM each), 0.25 μl Forward primer (100 pmol/μl), 0.25 μl Reverse primer (100 pmol/μl), 1 μl plasmid template (pJU204) from miniprep, 39 μl H₂O, 2 μl High fidelity PFU. The mixture was run with the PCR program as follows:

- **Step 1 – Denaturing**: 92°C for 2 mins
- **Step 2 – Denaturing**: 92°C for 0.5 min
- **Step 3 – Annealing**: 53°C for 0.5 min
- **Step 4 – Extension**: 68°C for 32 mins
- **12 cycles of Step 2 to Step 4**
- **Step 5 – Final extension**: 68°C for 10 mins

After the completion of the PCR reaction, the template (methylated DNA) was digested with DpnI. Then the PCR product was purified with Qiagen PCR purification kit and confirmed with agarose gel electrophoresis. The mutated plasmid was then transformed into *E. coli* (DH5α competent cells), grown, isolated, and confirmed with sequencing.

2.7 Yeast sporulation and tetrad dissection

Diploid yeast cells to be sporulated were streaked onto the sporulation plate. The cells were incubated at 30°C for 5-7 days. The cells were examined microscopically to determine tetrad formation. A loop of yeast colonies was resuspended gently into 100 μl H₂O, then 3 μl of β-glucuronidase (10 U/μl, Sigma #G-0751) was added and cells were incubated at RT for 25 mins. Using a pipettor, 20 μl of the cells were transferred to a tilted YPDA plate, allowing the droplet of cell suspension to run down the sloping agar surface. Then a loop was used to make streak lines of β-glucuronidase treated cells. Tetrads were dissected using a Singer MSM system micromanipulator. After dissection, the spores were incubated at 30°C for 4-5 days.
Mutations containing primers

Denature DNA template anneal mutagenic primers PCR

Primers design for point mutation introduction

Mutated plasmids with nicks

DpnI digestion

Transformation into E. coli, nick repair.

Primers design for sequence deletion

Figure 2.1 Site-directed mutagenesis by whole plasmid synthesis
2.8 Glycerol gradient sedimentation

Glycerol gradients were made and run essentially as described in Martin et al. (2002) and Bartels et al. (2003). 80 μl splicing extract was added to 120 μl GG buffer (20 mM HEPES (pH 7.0), 100 mM KCl, 0.2 mM EDTA) in order to dilute the glycerol concentration of the splicing extract from 20% to 8%. Then the diluted sample (200 μl) was layered onto 10-30% glycerol gradients (11 ml) containing GG buffer.

After sedimentation at 37000 rpm for 17 h in a SW40 Ti rotor (Beckman) at 4°C, fractions of 400 μl were collected and stored at -70°C. To analyse the snRNA distribution, collected fractions were deproteinised, then extracted with an equal volume of phenol chloroform and ethanol precipitated, then subjected to Northern analysis.

To analyse the protein distribution, the odd fractions were extracted by 400 μl phenol/chloroform/Isoamyl-alcohol. The aqueous phase was removed, the protein layer and phenol/chloroform/Isoamyl-alcohol was acetone precipitated by adding 1000 μl ice-cold acetone, mixed, and was placed at -20°C for 1 hr. The solution was spun at 14000g for 10 min at 4°C, and the supernate was removed. The protein pellets were washed 1 time with 1 ml cold 70% ethanol, and spun at 14000g for 5 min at 4°C. The supernate was removed and the pellet was dried in a speed vac for 10 mins. The pellet was resuspended in 15 μl 1X SDS loading buffer, and then subjected to SDS-PAGE and Western blotting.

2.9 Immunofluorescent staining

In order to compare the rna distribution profile in mutants and WT yeast strains, all tested yeast strains and their WT were grown, fixed, immunostained and viewed under the same light exposure. Cells were grown in 10 mls of YPDA to OD600 0.3 to 0.5 and fixed by the addition of 1/10 volume of 37% formaldehyde. The culture with fixative was kept in a 50 ml falcon tube and rotated for 40 mins at RT. Cells were then recovered by centrifugation at 3500 rpm for 5 mins, and washed gently with 1 ml of Buffer B (1.2M sorbitol, 100 mM KPO4, pH7.0) 3 times. To spheroplast, cells were resuspended in 500 μl of Buffer B containing 10 mM DTT, and 20 μl lyticase (20 mg/ml). The suspension was incubated at 30°C on a rotating wheel.

As soon as the settled cell pellet had lost its creamy, yellow colour and become translucent (about 30mins), cells were recovered, washed once with Buffer B and
resuspended in 100 μl of Buffer B. A drop of cell suspension was placed on a poly-L-lysine coated slide, allowed to sit for 20-30 min, and then excess solution was aspirated. (One additional step is needed for detection of tubulin: the cells were fixed with methanol for 5 min and subsequently fixed with acetone for 30 seconds at RT). PBST containing milk (5% w/v) was added to the samples, which were blocked for 20 min at RT. The cells were covered with PBST-5% milk containing the primary antibody.

The slide was set in a humidity chamber and incubated at 4°C overnight. Following this, excess solution was aspirated. The cells were washed with PBST twice for 20 min, and then cells were covered with PBST-5% milk containing secondary antibody. The slide was then set in a humidity chamber at RT for 1 hr, and excess solution was aspirated. The cells were washed 3 times with PBST for 20 min, and mounting solution (containing DAPI) was placed on the cells. A coverslip was placed on top and sealed with nail polish. Slides were stored at -20°C until examination.

2.10 Flourescent in situ hybridization

In order to compare the rna distribution profile in mutants and WT yeast strains, all tested yeast strains and their WT were grown, fixed, hybridised and viewed under the same light exposure. The protocol was adapted from the Singer lab protocol: In situ hybridization of yeast cells (http://www.singerlab.org/protocols/insitu_yeast.htm). Yeast cells were grown in 36 ml cultures in the appropriate media until reaching early log phase (OD₆₀₀ between 0.2-0.4). Cells were fixed for 10 mins at RT by directly adding to the medium 8 ml of 20% formaldehyde, 50% acetic acid. The fixative was removed by three rounds of washing in 10 ml of ice-cold buffer B (1.2 M sorbitol, 100 mM KPO₄, pH 7.0), followed by centrifugation (5 mins at 3500 rpm, 4°C).

Spheroplasting: cells were resuspended in 500 μl of Buffer B, 1 mM DTT, 20 μl lyticase (20mg/ml) and subsequently incubated at 30°C on a rotating wheel for 30 mins. Cells were then centrifuged for 2 mins at 3500 rpm at 4°C, and washed once with ice-cold buffer B. Cells were further resuspended in 500 μl of buffer B, and 100 μl was added to a poly-L-lysine coated slide. Cells were then left to adhere to the slide by incubating them 30 mins at 4°C. Slides were washed once with ice-cold buffer B, and then stored in 70% ethanol at -20°C (the slides can be stored for weeks).

Hybridization: cells were rehydrated for 5 mins at RT, in 2x SSC (300 mM NaCl, 30 mM sodium citrate, pH 7.0), 50% formamide. Cells were hybridized overnight in 40 μl
(hybridisation buffer) of a mixture containing 10% dextran sulfate, 0.02% RNase-free BSA, 40 μg E. coli tRNA, 2 x SSC, 50% formamide, 1 μl RNAsin and 1 μl 100 pmol/μl probe (Hybridisation temperature: 37°C for U2, U3, and U6 probes, 42°C for U5 probe). For U5 and U6 probes, additional 0.1% NP40 (v/v) is added in hybridisation buffer. Then the cells were washed twice for 20 mins in 3x SSC, 50% formamide in 37°C (U2, U3, and U6 probes) or 42°C (U5 probe). Cells were then mounted with vectashield containing DAPI.

2.11 Prp8p (8.6) antibody production

8.6 antibodies against the N-terminal peptide of yeast Prp8p were first raised by A. Gordon, Beggs lab. The sequence, apart from the first cysteine residue, corresponded to aa2-35 of Prp8p. The purpose of adding the cysteine that contained a free sulphhydryl group was for peptide conjugation. The peptides were conjugated with PPD (purified protein derivative of *Mycobacterium tuberculosis*), and subsequently were used to produce the 8.6 antibodies (batch name: 666). As the amount of 666 antibodies was low in the Beggs lab, I decided to re-raise the 8.6 antibodies.

The peptides were synthesised by Sigma Genosys Ltd., but conjugated with KLH (Keyhole Limpet Hemocyanin). This was because the PPD was no longer commercially available. From Diagnostics Scotland, I obtained pre-immune serum from ten rabbits and tested for antigenicity to yeast crude extracts. Two rabbits (R1689 and R1703) that showed low reaction with yeast proteins were chosen to raise the antibodies, and the injection of antigen and bleeding of the rabbits were performed by Diagnostics Scotland. In this thesis, only serum from rabbit R1689 was used.
Chapter 3

Prp8p Partitioned Reveals Protein-protein Interactions

3.1 Introduction

A genetic screen by Andrew Newman's group (Cambridge) showed that Prp8p could be partitioned into several domains using a custom-built Tn5-based transposon (Figure 3.1A). This system allows expression of Prp8p as two contiguous polypeptides (Figure 3.1B), capable of trans-complementation and supporting the cell's viability. Thus, four functional domains can be defined in the 2413 residue Prp8p, with boundaries in the regions 416-436, 770-773 and 2170-2173. The central region of the protein has so far resisted dissection by this approach. The partitioned constructs allowed investigation of factors that associate with the N- or C-terminal regions of Prp8p by isolation of the sub-complexes associated with different Prp8 protein fragments. Thus, the U5 snRNP protein Snu114p associates with Prp8p in the region 437-770, whereas fragmenting Prp8p at residue 2173 destabilises its association with Aar2p.

3.2 Partitioned Prp8p

Strains with PRP8 partitioned at aa770, and aa2173 were constructed on the genome by insertion of pAW28 through homologous recombination (Figure 3.1C). The clones that survive on −Ura dropout plates were tested by yeast colony PCR. For PRP8 partitioned at aa436, the PCR insertion fragment was generated from plasmid M529, which encoded prp8 with insertion of the transposon at aa436 (Figure 3.1B). PRP8 partitioned at aa436 construct was expressed as two polypeptides (aa1-439, aa436-2413), which contained three amino acids repeat that are essential for cell viability (Section 3.4). The aa436 partition clones were selected under G418 geneticin, and tested by yeast colony PCR. Subsequently, the size of Prp8p C-terminal fragments produced in WT and partition strains were confirmed by Western blotting (Figure 3.2A).
A) The transposon was custom-built from 3 main components: 1) a fragment of the PRP8 gene 3' UTR containing signals for the transcription termination and poly-adenylation and stop codons in all 3 reading frames; 2) KanMX4 cassette (Wach et al., 1994) conferring Kanamycin-resistance in E. coli and G418-resistance in yeast; 3) a fragment of the PRP8 gene 5' UTR containing signals for initiation of transcription and translation. The transposon is flanked by 19 bp 'Mosaic End' sequences (CTGTCTCTTATACACATCT) recognised by Tn5 transposase (ME).

B) Formation, structure and expression of a dissected prp8 gene resulting from insertion of a transposon into the open reading frame (not to scale). The transposon terminates transcription of the upstream prp8 gene sequences and drives expression of the prp8 gene sequence downstream. Prp8p is therefore expressed as two contiguous protein fragments via translation of two separate mRNAs. Each fragment is tagged at its N-terminal with a SV40 Nuclear Localisation Signal (NLS) to ensure targeting to the nuclear compartment. In the case of partitioned Prp8p after codon 436, transcription introduced a 9 bp Direct Repeat (DR) of the target sequence flanking the transposon (upper section of diagram). For partition after codon 770 and 2173, the transposon was generated in such a way as to avoid the 9 bp repeat (lower section of diagram).

Figure 3.1 Structure of the gene dissection transposon and partitioned Prp8p

A) The transposon was custom-built from 3 main components: 1) a fragment of the PRP8 gene 3' UTR containing signals for the transcription termination and poly-adenylation and stop codons in all 3 reading frames; 2) KanMX4 cassette (Wach et al., 1994) conferring Kanamycin-resistance in E. coli and G418-resistance in yeast; 3) a fragment of the PRP8 gene 5' UTR containing signals for initiation of transcription and translation. The transposon is flanked by 19 bp 'Mosaic End' sequences (CTGTCTCTTATACACATCT) recognised by Tn5 transposase (ME).

B) Formation, structure and expression of a dissected prp8 gene resulting from insertion of a transposon into the open reading frame (not to scale). The transposon terminates transcription of the upstream prp8 gene sequences and drives expression of the prp8 gene sequence downstream. Prp8p is therefore expressed as two contiguous protein fragments via translation of two separate mRNAs. Each fragment is tagged at its N-terminal with a SV40 Nuclear Localisation Signal (NLS) to ensure targeting to the nuclear compartment. In the case of partitioned Prp8p after codon 436, transcription introduced a 9 bp Direct Repeat (DR) of the target sequence flanking the transposon (upper section of diagram). For partition after codon 770 and 2173, the transposon was generated in such a way as to avoid the 9 bp repeat (lower section of diagram).
Figure 3.2.

A) Western blotting to detect the Prp8p C-terminal polypeptides. Prp8p C-terminal fragments were first pulled down by horse IgG agarose and then fractionated in a 4-12% PAGE gel, blotted and probed with Peroxidase anti-peroxidase (PAP) antibody. Lane 1) RG8T, 2) BMA38a, 3) KL8T/H436, 4) KL8T/770, 5) RG8T/2173.

B) Prp8p WT and Prp8p partition strains were plated on YPDA or YPGR plates with initial of cells OD<sub>600</sub> 0.4, and growth at 18°C, 23°C, 30°C, and 37°C for 7 days before being photographed. KL8T/770, KL8H/436, and KL8H/436T were shown heat sensitive.
3.3 Partition strain growth phenotypes

I investigated the effect of Prp8p partitions on the strains. The WT strain BMA38a (Albers et al., 2003), RG8T (full length Prp8p C-terminal TAP tag), RG8T/2173 (Prp8p C-terminal TAP tag and partitioned at aa2173), KL8T/770 (Prp8p C-terminal TAP tag and partitioned at aa770), KL8H (Prp8p full length with C-terminal triple HA tags), KL8H/436 (Prp8p C-terminal triple HA tags and partitioned at aa436), KL8H/436T (Prp8p C-terminal triple HA tags, partitioned at aa436, and TAP tag at the C-terminal of aa1-436 fragment) were grown in YPDA broth and plated on YPDA plates in serial dilutions. Then the plates were incubated at different temperatures; 18, 23, 30, 37°C. BMA38a, RG8T, RG8T/2173, KL8H showed growth at all the temperatures, but KL8T/770, KL8H/436, and KL8H/436T were heat sensitive (Figure 3.2A). All the Prp8p partitioned strains showed slow growth when compared to un-partitioned Prp8p strains.

I have tried to create a Prp8p aa436 partition strain in host RG8T, but I failed to obtain any positive clones by insertion of the transposon generated from pAW28 (refer to Section 3.4). This could be due to the Prp8p C-terminal TAP-tag, as the TAP-tag may disrupt the function of Prp8p when combined with the aa436 breakpoint. However, the insertion of the transposon into Prp8p aa436 was achieved with no difficulty in strain BH8T (PRP8:TAP-TRP1, HIS3-GAL1-3HA:PRP8, otherwise as BMA38a). This is possibly due to over-expression of aa1-439 Prp8p stabilising the association of Prp8p aa436-2413 TAP-tag polypeptides. This strain showed a better growth rate compared to KL8H/436, and KL8H/436T, but this construct did not rescue the heat sensitive phenotype caused by Prp8p aa436 partitioning (Figure 3.2B)

3.4 Requirement for 3 amino acids repeats and Mosaic End sequences for cell viability in the Prp8p aa436 partition strain

For Prp8p partitioning at the aa436 region, a total of 796 yeast colonies transformed by insertion of PCR product from pAW28 that grew on –Ura dropout plates were screened. All the colonies that were screened yielded PCR product from non-insertion regions or both the non-insertion and insertion regions. Western blotting identified the size of Prp8p in these clones, which showed no reduced size of Prp8p in the clones tested. Apparently the clones were false positives, and the presence of both positive and negative insertion PCR products could be due to the duplication of genomic PRP8 that caused by mutation.
Figure 3.3
A) Tetrad dissection of clone derived from BMA38 (2n) with transposon inserted at PRP8 aa436. B) Yeast colony PCRs for URA3 insertion confirmation over PRP8 in BMA38 (2n) transformants and its haploid cells. (U: ycPCR from upstream insertion region to URA3, y: ycPCR before and after the insertion). C) The transposon inserted diploid cells and its surviving haploid cells were plated on YPDA and –Ura dropout plates. BMA38 (2n) without transposon insertion served as negative control and transposon inserted BMA38 (2n) cells served as positive control over the surviving haploid cells after the tetrad dissection.

As partitioning of Prp8p at the aa436 region by inserting pAW28 construct might cause cell lethality, transformation was performed on diploid BMA38. Clones showing positive insertion of URA3 gene after yeast colony PCR screening were sporulated, and tetrads were dissected. Two clones were picked for the sporulation and tetrad dissection.
tests. Results showed only two surviving spores after the tetrad dissections (Figure 3.3A). The surviving colonies were tested by yeast colony PCR (ycPCR) for URA3 insertion (Figure 3.3B) and plated on both YPDA and -Ura dropout plates (Figure 3.3C). Results show that URA3 is undetected in ycPCRs, and the failed of grown of colonies plated on -Ura dropout plates confirmed this. These indicate that partitioning of PRP8 at the 436 aa region by insertion of pAW28 transposon PCR product is lethal to the cell.

Partitioning of Prp8p at the aa436 region caused cell death was rather surprising. Our collaborator, Dr. Andrew Newman obtained a collection of clones, in which Prp8p was partitioned at region aa416-436 by using the transposon techniques. One difference between our partitioned strains is that, Dr. Andrew Newman Prp8p aa416-436 region partition clones contained 3 amino acids (aa) repeats (for example: Prp8p aa1-439, aa437-2413), and Mosaic End sequences.

A plasmid (M529) that contained prp8 with transposon insertion at aa436 was obtained from Dr. Andrew Newman. Transformation PCR product was generated by using M529 as template. Before the transformation, the PCR product was digested with DpnI in order to digest the methylated PCR template (M529) that could potentially cause false positive insertion. By using this strategy, I succeeded in obtaining clones with Prp8p partitioned at aa436: KL8TG/436, KL8H/436, and KL8H/436T.

3.5 TAP-tag purification of Prp8p complexes

To determine the extent to which Prp8p complexes were intact or were disrupted by the dissection, extracts from parent (RG8T) and dissected strains, KL8T/770 and RG8T/2173, were subjected to the TAP affinity-selection procedure and the co-selected proteins were separated by SDS-PAGE and analysed by mass spectrometry. Proteins that were co-selected with full-length Prp8p included Aar2p, Brr2p, Prp3p, Prp4p, Prp6p, Snu114p and several Sm and Lsm proteins (Figure 3.4A). These represent a subset of the proteins that were previously shown to be components of the U5 snRNP or U4/U6.U5 tri-snRNP. The Prp8T/770p C-terminal fragment (Prp8p C-terminal TAP-tagged and partitioned at aa770) co-selected Aar2p, Brr2p, Prp3p, Prp4p, Prp6p, and Spp382p. The Prp8T/2173p C-terminal (2174-2413) co-selected a protein similar in size to Brr2p, whereas other co-selected proteins were in very low abundance due to the reduced stability of the complex.
Figure 3.4 TAP purification of Prp8p-associated proteins

A) TAP-purified proteins were fractionated in a 4-12% PAGE gel, stained overnight with Sypro ruby protein stain and analysed by mass spectrometry. B) TAP-purified proteins associated with Prp8p dissected residue 2173 (lane 1) or 770 (lane 2) were fractionated in a 4-12% PAGE gel, blotted and probed with anti-8.6 antibodies. C) Effect of salt on the association of the pairs of dissected Prp8 polypeptides. Splicing extracts were incubated with IgG-agarose in the presence of 150, 300 or 500 mM NaCl. The pull-down proteins were fractionated by 4-12% PAGE gel, blotted and probed with anti-8.6 antibodies. Lane 1) Non-tagged full-length Prp8p, 2) Full-length Prp8p with TAP tag at the C-terminus, 3) TAP-tagged Prp8p dissected at residue 770, 4) TAP-tagged Prp8p dissected at residue 2173. Proteins were fractionated, blotted and probed with anti-8.6 antibodies.

Interestingly, Snu114p, which was co-selected efficiently with full-length Prp8p, was hardly detectable in association with Prp8T/770p. In contrast, the amount of Aar2p associated with Prp8T/770p was increased compared to the full-length Prp8p. Aar2p was
previously found associated with a 16S form of yeast U5 snRNP that contained Prp8p, Snu114p and the Sm proteins but lacked the majority of known Prp8-associated proteins, including Brr2p, and it was suggested that this Aar2-U5 snRNP complex represents an intermediate particle in U5 snRNP biogenesis. In addition, Spp382p was associated with Prp8T/770p, but was not detectably associated with full-length Prp8p or with Prp8T/2173p. Spp382p has been affinity-selected previously with Prp8p and other splicing factors (Hazbun et al., 2003; Gavin et al., 2002), but is not known to be a stable component of U5 snRNPs. Thus, Prp8T/770p produced a different profile of co-selected proteins compared to full-length Prp8p. This presumably represents more than one Prp8p complex, and likely includes the Aar2p-U5 snRNP precursor complex related to that described by Gottschalk et al. (2001). The profile of TAP-purified proteins visualised by Sypro ruby staining (Figure 3.4a) suggested that this procedure efficiently separated the N-terminal and C-terminal fragments of Prp8p dissected at positions 770 and 2173. Nevertheless, using Western blotting with anti-8.6 antibodies specific for the N-terminus, the N-terminal fragments (1-770 and 1-2173) were detectable in the TAP-purified fraction, albeit at a low level, as shown in Figure 3.4B.

In order to establish conditions that gave complete separation of the N-terminal and C-terminal Prp8p fragments, the effects of different salt conditions on the association of Prp8p N- and C-terminal polypeptides were tested in a single-step precipitation. The Prp8p C-terminal TAP-tagged fragment was affinity-selected using IgG agarose beads, and the presence of the Prp8p N-terminal fragment in co-precipitates was investigated by Western blotting with anti-8.6 antibodies (which recognise the N-terminal fragments of Prp8p but also interact with the protein A moiety of the TAP-tag on the C-terminal fragments). This single-step affinity-selection in the presence of 150 mM or 300 mM salt did not disrupt the association of the N- and C-terminal Prp8p fragments, however, with 500 mM salt, only C-terminal TAP-tagged Prp8p polypeptides were selected (Figure 3.4C).

3.6 Co-immunoprecipitation of Snu114p and Aar2p with Prp8p

As the Prp8p polypeptide pairs can be dissociated in 500 mM NaCl, the association of different fragments of Prp8p with other splicing factors was investigated in this condition. The TAP-tagged C-terminal Prp8p fragments were pulled down with IgG-agarose, the supernates were then incubated with anti-8.6 antibodies to pull down the N-terminal of Prp8 polypeptides (Figure 3.4A), and these immune precipitates were analysed by Western
Figure 3.5 Detection of proteins associated with the N- or C-terminal fragments of Prp8p. A) Splicing extracts were first incubated with horse IgG agarose in 500 mM NaCl to pull-down only the C-terminal Prp8p fragments (lanes 1 to 4). The supernates were then transferred into new tubes and incubated with anti-8.6-adsorbed protein A Sepharose beads in the presence of 500 mM NaCl (lanes 6 to 8). After PAGE, Prp8p fragments that bound to the beads were detected by Western blotting, using anti-8.6 antibodies. Lanes 1 & 6) full-length TAP-Prp8 protein; lanes 2 & 7) full-length non-tagged Prp8p; lanes 3 & 8) Prp8T/436p; lanes 4 & 9) Prp8T/770p; lanes 5 & 10) Prp8T/2413p. The blot from a) was stripped and reprobed B) with anti-Snu114 antibodies or C) with anti-Aar2p antibodies. D) Splicing extracts from strains producing non-tagged full-length Prp8p (lane 1), Prp8T/770p (lanes 2-4), Prp8T/2173p (lanes 5-7), and full-length Prp8Tp (lanes 8-10) were incubated with IgG-agarose in the presence of 150 mM NaCl. The precipitates (Co-IP), 10% total, and 10% supernates (Sup) after the pull-down (only precipitate from non-tagged Prp8p) were analysed by Western blotting with both anti-8.6 and anti-Aar2p antibodies.
blotting for the presence of Prp8p N-terminal fragments (Figure 3.5A), Snu114p (Figure 3.5B) or Aar2p (Figure 3.5C). Snu114p was detected associated with TAP-tagged full-length Prp8p (Figure 3.5B, lane 1) and with C-terminal Prp8p polypeptide 437-2413 (Figure 3.5B, lane 3), but not with Prp8p 771-2413, Prp8p 2174-2413 or with full-length untagged Prp8p that served as a control, (Figure 3.5B, lanes 4, 5 and 2). Following the precipitation with anti-8.6 antibodies, Snu114p was found associated with non-TAP-tagged Prp8p (Figure 3.5B, lane 7), and with N-terminal polypeptides 1-770 and 1-2173 of Prp8p (Figure 3.5B, lanes 9 & 10). The amount of Snu114p pulled down with full-length TAP-tagged Prp8p, and with N-terminal amino acids 1-439 of Prp8p was low (Figure 3.5B, lanes 6 & 8), as most Snu114p had already been pulled out of these samples during the first-round precipitation of the TAP-tagged C-terminal polypeptides (Figure 3.5B, lanes 1 & 3). Thus, Snu114p associates with the region 437-770 of Prp8p.

The same blot was stripped and re-probed with anti-Aar2p antibodies. Aar2p was present in precipitates with TAP-tagged, full-length Prp8p, and with Prp8p residues 437-2413, and 771-2413 (Figure 3.5C, lanes 1, 3, & 4), but not with untagged Prp8p, or with Prp8p amino acids 2174-2413 (Figure 3.5C, lane 2 & 5). Unfortunately, the presence of Aar2p in the N-terminal Prp8p pull-downs was obscured by the presence of antibodies. However, even when the Prp8T/2173p (Prp8p C-terminal TAP-tagged and partitioned at aa2173) immunoprecipitation was performed under conditions (150 mM NaCl) that did not disrupt the association of the N- and C-terminal Prp8p fragments, Aar2p was not coprecipitated (Figure 3.5D, lane 7).

In addition, centrifugation through a 10-30% glycerol gradient (100 mM KCl) showed that Aar2p mostly co-fractionates with full-length Prp8p and Snu114p (Figure 3.6A, lanes 5-11), although some Aar2p was also detected near the top of the gradient (Figure 3.6A, lane 25) as was observed previously by Gottschalk et al. (2001). In contrast, the Aar2p in Prp8T/2173 extract was mostly found in the lower density fraction (Figure 3.6B, lane 25) and only a trace amount was detected in a higher density fraction (Figure 3.6B, lane 7), after the film was over-exposed. Thus, dissecting of Prp8p at position 2173 disrupted the interaction between Aar2p and Prp8p.
Figure 3.6
Glycerol gradient fractionation of Prp8p, Snu114p and Aar2p. RG8T cell extract containing full-length Prp8p-TAP (A) and RG8T/2173p extract containing Prp8T/2173p (B) were fractionated in 10-30% glycerol gradients. Alternate gradient fractions were acetone precipitated and proteins were analysed by Western blotting with anti-Snu114p and anti-Aar2p antibodies. Asterisk (*) indicates the presence of Aar2p near the top of the gradient (Figure 3.6A and B, lane 25).
3.7 Discussion

Yeast Prp8p is highly conserved in both its sequence and its size, being 61% homologous to human Prp8p. Besides the MPN (Mpr-1, Pad-1, N-terminal; residues 2173-2310) and RRM (RNA Recognition Motif; 1059-1151) domains that were predicted by bioinformatic approaches, and a functionally proven nuclear localisation signal (Chapter 4), there is no other obvious structural domain within this huge protein. By functional, genetic classification, Brow and colleagues (Kuhn et al., 2002) designated 5 domains (Region a: residues 233-365, Region b: residues 608-687, Region c: residues 785-864, Region d: residues 1091-1197, and Region e: residues 1621-1878).

The results presented here show that Prp8p can be dissected into four functional domains: residues 1-439, 437-770, 771-2173, and 2174-2413. The dissection points near the C-terminus of Prp8p (2170-2179) lie precisely at one boundary of the predicted MPN/JAB domain (2173-2310). This is consistent with the notion that these transposon insertions may be tolerated because they have targeted a ‘linker’ sequence upstream of this functional module. Furthermore, the Prp8p NLS, and genetically defined Region a are within residues 1-439, Region b is within 437-770, and the RRM and Regions c, d, e are within 771-2173. The failure to obtain viable constructs with transposon insertions in the 771-2173 region of Prp8p may indicate that Regions c, d, e and the RRM function as one indivisible unit.

The combination of Prp8p partition and TAP affinity-selection were used to reveal the potential regions of protein interaction between Prp8p and other splicing factors. By increasing the salt concentration to 500 mM, the associations of the N- and C-terminal Prp8p polypeptides were disrupted. Therefore, using a two-single-step immunoprecipitation strategy (Figure 3.5), the association of Snu114p with Prp8p was found to require Prp8p residues 437-770. This in close agreement with two-hybrid interaction data in which Snu114p as bait selected multiple Prp8p fragments that had amino acids 420 to 464 in common (Dix and Beggs, unpublished results). The strong association of Prp8p and Snu114p is not surprising, as Achsel et al., (1998) showed that human Prp8 and Snu114 proteins remained associated in 0.4 M of the chaotropic salt sodium thiocyanate.

Interestingly, even at moderate salt concentrations (150 mM NaCl or 100 mM KCl) the association of Aar2p was disrupted by dissecting Prp8p at residue 2173. Previously, Gottschalk et al. (2001) purified yeast Aar2p associated with Prp8p and Snu114p. Here I show that Prp8p/771-2413p could co-select Aar2p in the absence of Snu114p (Figure 3.5C, lane 4), and Prp8T/2173p could co-select Snu114p in the absence of Aar2p (Figure 3.5B, lane 10 and Figure 3.5B). This suggests that the associations of Snu114p and Aar2p with
Prp8p are not interdependent.

Although Prp8p has been dissected into pairs of functional polypeptides in this work, each of these disruptions was detrimental to growth; the 2173 dissection caused slow growth and the 436 and 770 dissections resulted in heat sensitivity. For the 2173 dissection at least, the integrity of Prp8p's interaction with another protein was apparently reduced stability of the complex. This confirms the functional significance of the very large size of this highly conserved protein that has been proposed to function as a scaffold for molecular interactions in the spliceosome.
Chapter 4

U5 snRNP assembly, from cytoplasm to nucleus

4.1 Introduction

In yeast, the function of the UsnRNP-specific proteins in the completion of the biogenesis of U snRNPs is still poorly understood. In this study, I have identified the nuclear localisation signal (NLS) domain of Prp8p by sequence alignments and confirmed it by biochemical test. Subsequently, the role of Prp8p in U5 snRNP biogenesis, and the effect of deleting the Prp8p NLS on other U5 snRNP protein localisations and association were investigated. This is the first report that proposes that yeast U5 snRNA travels to the cytoplasm, where it forms a Prp8p-Snu114p-Aar2p U5 complex, before re-entering to the nucleus and forming the mature U5 snRNP with Brr2p and other U5-specific proteins. The C-terminal region of Prp8p is important for Brr2p association (van Nues and Beggs, 2001), and mutations resembling those in human Prp8p that cause retinitis pigmentosa type 13 decreased the association between Prp8p and Brr2p. Furthermore, the importin Kap95p has been identified to be essential for Snu114p, and snRNA (U2, U5, and U6) nuclear import.

4.2 Prp8p NLS identification

By using the ProteinPredict server (Rost et al., 2003), a nuclear localisation signal (NLS) was identified at aa96-117 of yeast Prp8p. To validate the prediction, firstly I investigated the localisation of Prp8p with aa96-117 deleted compared to WT Prp8p. In the second step, I looked into the localisation of GFP fused to Prp8p aa96-117 compared to GFP alone.

To delete aa96-117 from the full-length Prp8 protein, I employed site-directed mutagenesis (as mentioned in section 2.6) on plasmid pJU204 (CEN-HIS3; a plasmid that harbours full-length PRP8 with its own promoter and encoding a triple-HA tag at the C-terminus of the protein) (Umen and Guthrie, 1995). The deletion of aa96-117 in the pJU204 sequence was confirmed by sequencing. Full-length pJU204 and pJU204ΔNLS were transformed into KL1 (yeast strain with genomic PRP8 under regulation of P_GAL1), and selected on -His YMGR plates. Both strains that contained pJU204 and pJU204ΔNLS were grown in YPDA broth and the localisation of Prp8p was investigated by immunofluorescent
staining. The immunofluorescent staining to detect the triple-HA tag of plasmid-based Prp8p showed that the full length Prp8p was localised in the nucleus (Figure 4.1A). However, the Prp8p with deletion Δaa96-117 was present in both nucleus and cytoplasm (Figure 4.1B).

A plasmid (pGFP-N-NLS) that encodes aa96-117 fused to GFP was constructed from PGFP-N-FUS. PGFP-N-FUS (CEN-URA3-P_{MET25}-GFP) is a plasmid that expresses GFP under control of MET25 promoter (Niedenthal et al., 1996). Both the pGFP-N-FUS and pGFP-N-NLS were transformed into yeast cells separately and selected on YMM (-Met, -Ura) dropout plates. The transformed clones were then grown on YMM (-Met, -Ura) broth until log-phase, and subsequently the cells were harvested and examined microscopically. GFP without insertion was distributed throughout the cell (Figure 4.1C). However, when GFP was fused with aa96-117 of Prp8p, the fluorescence was concentrated in the nucleus (Figure 4.1D). This confirmed that aa96-117 of yeast Prp8p encodes a functional NLS.

![Figure 4.1. Localisation of Prp8p. A) Full length Prp8p B) Deletion-NLS-Prp8p. The full length Prp8p was sub-nuclear localised, but the deletion-NLS-Prp8p was detected in both the nucleus and cytoplasm. C) Localisation of GFP and D) GFP fused to the Prp8p NLS (aa96-117 Prp8p). GFP was found distributed throughout the cells, but concentrated in the nucleus when fused to aa96-117 of Prp8p.](image)
One interesting finding was that KL1 strain that contained ΔNLS-Prp8p grew slowly on YPDA plates, but survived. In addition, the strain showed no temperature-sensitive or cold-sensitive phenotype (Figure 4.2A). The efficiency of splicing \textit{in vitro} from KL1-pJU204ΔNLS was reduced, but no step 1 or 2 defect was observed when compared to the strain KL1-pJU204 (Figure 4.2B). Another interesting finding was that, ΔNLS Prp8p immunoprecipitated reduced levels of U4 and U6 snRNAs, but comparable levels of U5 to the full-length Prp8p immunoprecipitates (Figure 4.2C), despite the fact that a portion of deleted-NLS-Prp8p is located in the cytoplasm.

\textbf{Figure 4.2. The effect of deletion-NLS-Prp8p on cells.} A) The plasmid (pJU204) containing full length \textit{PRP8} and \textit{deletion-NLS-prp8} constructs were transformed into KL1 and plated on -His dropout plates. The transformants were restreaked on YPDA and incubated at different temperatures (14°C, 23°C, 30°C, and 37°C). Cells containing deletion-NLS-Prp8p showed slow growth compared to cells with full-length Prp8p, but it has neither TS nor CS phenotype. B) \textit{In vitro} splicing was performed on extracts derived from cells with full length Prp8p or deletion-NLS-Prp8p. Deletion-NLS-Prp8p extract showed a reduced rate of splicing \textit{in vitro} when compared to full length Prp8p extract, but no accumulation of step I or step II products was observed. C) Immunoprecipitation of Prp8p showed that deletion-NLS-Prp8p precipitated less U4 and U6 snRNAs, but comparable amounts of U5 snRNA when compared to full length Prp8p precipitation.
A more surprising finding is that aa96-117 of yeast Prp8p shows low homology to Prp8ps from other species (Figure 4.3). This suggests that the aa96-117 sequences could be exclusively present in budding yeast. By using PSORTII (Nakai and Horton, 1999), Grainger and Beggs (2005) showed a potential bi-partite NLS is located at aa96-156. This indicates that the additional NLS that located at aa118-156 of yeast Prp8p might responsible for the nuclear import of the deletion aa96-117 Prp8p. However, whether the aa118-156 of yeast and other eukaryotic Prp8ps serve as a functional NLS is yet to be confirmed.

Figure 4.3. Alignments of N-terminal of Prp8 orthologs.
All sequences aligned using ClustalW to aligned the N-terminal regions. The polyproline tract of S. cerevisiae is underlined and the NLS sequence is highlighted. (numbers refer to the amino acid residues).
4.3 Effect of Prp8p delocalisation on localisation of other U5 components

The 16S and 20S forms of the U5 snRNP were identified by the Luhrmann and Abelson groups. The 16S U5 snRNP was isolated by pulling down Aar2p (Gottschalk *et al.*, 2001), and consists of Prp8p, Snu114p, Sm proteins, Aar2p, and U5 snRNA. The 20S U5 snRNP was isolated by using Brr2p as bait, and consists of every component of 16S U5 snRNP, except Aar2p, and also includes Prp28p, Snu40p, and Dib1p (Stevens *et al.*, 2001).

To investigate the effect of ΔNLS Prp8p on localisation of U5 snRNP components, plasmids encoding full-length Prp8p and ΔNLS-Prp8p were transformed separately into KL1-Aar2 (KL1 with Aar2p TAP tag), KL1-Brr2, KL1-Snu114, KL1-Lsm8, KL3-SmB, KL3-SmD1, and KL3-SmE yeast strains. In both KL1 and KL3 yeast strains, genomic PRP8 is regulated by the PGAL1 promoter. After that, the yeast cells that harboured the plasmid were grown in YPDA broth. Subsequently, TAP-tagged proteins were immunofluorescent stained and observed microscopically as mentioned in the materials and methods (2.9).

Firstly, the localisation of the common components in both 16S and 20S U5 snRNPs were investigated. The results showed that Snu114p is in the nucleus in the WT strain (Figure 4.4A) and was delocalised to the cytoplasm in the ΔNLS Prp8p strain (Figure 4.4B). SmB1p, and SmD1p (that contain their own NLS (Bordonne, 2000)) are consistently in the nucleus in both WT and the ΔNLS Prp8p strains (Figure 4.4E, F, G & H). SmE1p showed a slight increase in the cytoplasm (Figure 4.4J) in the ΔNLS Prp8p strain when compared to the WT Prp8p strain (Figure 4.4I). Strikingly, U5 showed increased signal in the cytoplasm in the ΔNLS Prp8p strain (Figure 4.4P), when U5 was mainly detected in the nucleus in the WT Prp8p strain (Figure 4.4O).

Aar2p, that is present in 16S U5 snRNP, was found in both cytoplasm and nucleus in the WT Prp8p strain (Figure 4.4C). In the ΔNLS Prp8p strain, Aar2p appeared to be slightly increased in the cytoplasm, but this may not be significant (Figure 4.4D). Brr2p (present in 20S U5 snRNP) (Figure 4.4K & L), and other non-U5 snRNP components: U2 (Figure 4.4M & N), U6 (Figure 4.4Q & R), and Lsm8p (Figure 4.4S & T) were found to be consistently nuclear regardless of the localisation of Prp8p.

To further analyse the level of U5 that is bound to Aar2p associated complex, cell extracts derived from KL1-Aar2-pJU204 and KL1-Aar2-pJU204ΔNLS grown in glucose medium were utilised. Results showed that Aar2p in ΔNLS Prp8p yeast strain precipitated increased levels of U5 (S & L) when compared to WT Prp8p yeast strain (Figure 4.5A). This suggests an increased level of 16S U5 snRNP when Prp8p is delocalised to the cytoplasm.
Figure 4.4. The effect of deletion-NLS-Prp8p on localisation of other splicing components. The splicing factors under investigation were either TAP-tagged or protein-A-tagged in yeast strains (KL1 or KL3) with genomic PRP8 under PGAL promoter. Plasmid encoding full length Prp8p or deletion-NLS-Prp8p construct were transformed into genomic-PRP8-regulated cells and selected on -His dropout plates. The transformants were grown in YPDA broth and probed with rabbit anti-Protein A primary antibodies, then probed with anti-rabbit antibodies conjugated with Alexa fluor 594 dye. For snRNA localisation analysis, KL1 containing either full length Prp8p or deletion-NLS-Prp8p was grown in YPDA and fluorescent in situ hybridised with Cy3-labeled probe. A) Snu114p was shown to be nuclear localised in full length Prp8p background, but delocalised to the cytoplasm when the Prp8p NLS was deleted (B). SmE1p (J) and U5snRNA (P) showed increased cytoplasmic signals in deletion-NLS-Prp8p background. C & D) Aar2p was present in both cytoplasm and nucleus. SmB1p (E & F), SmD1p (G & H), Brr2p (K & I), U2 (M & N), U6 (Q & R), and U6 core protein Lsm8p (S & T) are nuclear in both full length Prp8p or deletion-NLS-Prp8p cells.
4.4 Prp8p mediates Aar2p nuclear localisation

Aar2p is present in both nucleus and cytoplasm when Prp8p is endogenously expressed (Figure 4.4C). To investigate the localisation of Aar2p in relation to the expression level of Prp8p, strain KL1-Aar2 was utilised. KL1-Aar2 was derived from BMA38a, with Aar2p-TAP and PRP8 regulated under the $P_{GAL}$ promoter. When Prp8p was over-expressed (in galactose medium), Aar2p was observed to be nuclear localised (Figure 4.5B, upper panel). When Prp8p was depleted by growth in glucose medium for 12 hours, the localisation of Aar2p was observed in both cytoplasm and nucleus (Figure 4.5B, lower panel), resembling the observation in strain expressing endogenous Prp8p (Figure 4.4I).

**Figure 4.5.** A) Immunoprecipitation of Aar2p from splicing extracts derived from full length Prp8p or deletion-NLS-Prp8p. Increased levels of U5 snRNA were precipitated by Aar2p in a deletion-NLS-Prp8p cell extract. B) Yeast strain KL1, with genomic PRP8 under $P_{GAL}$ promoter and Aar2p TAP tagged, was first grown in YPGR (galactose and raffinose medium) and the Aar2p was detected by rabbit anti-Protein A antibodies, followed by Alexa fluor-anti-rabbit secondary antibodies (upper panel). The cells were then shifted into YPDA (glucose medium) for 12 h, and Aar2p was detected throughout the cells (lower panel).
4.5 Retinitis pigmentosa type 13 in yeast

By using the published results from McKie et al. (2002), the residues in yeast Prp8p according to the mutations in human Prp8p that cause retinitis pigmentosa type 13 were mutated (Figure 4.6A). By using pJU204 (plasmid encoding full-length Prp8p), three separate prp8 point mutation plasmids were created; R2388K (rp5), R2388G (rp6), and F2392L (rp7), and confirmed by DNA sequencing.

The wild-type and mutated plasmids were then transformed separately into KL3-Brr2 (a yeast strain with genomic PRP8 regulated by the P_{GAL} promoter and with Brr2p C-terminally tagged with 13-Myc tag). Transformed yeast cells were first selected on a -His dropout plate, and subsequently plated on YPDA. The plated yeast cells were grown at 14°C, 23°C, 30°C, and 37°C. Wild type and rp7 showed growth at all the temperatures, but rp5 and rp6 were heat sensitive (Figure 4.6B). All the rp mutation strains showed slow growth when compared to wild-type Prp8p strains. The localisation of the mutated Prp8p was analysed by immunoflourescent staining. The yeast strains were grown in YPDA broth at 30°C. The localisations of these mutated Prp8ps were consistently nuclear (Figure 4.6C), resembling the localisation of wild-type Prp8p (Figure 4.1A). However, when rp5 was grown at 37°C, Prp8p was detected in both the nucleus and cytoplasm (Figure 4.9C).

4.6 Prp8p, Aar2p, Brr2p and U5 snRNA associations

To facilitate the analysis of associations of Aar2p, Snu114p, Brr2p, and Prp8p, the yeast strain KL3-Brr2 was used, in which Brr2p is tagged at the C-terminal with 13-Myc. Splicing extracts derived from this strain (KL3-Brr2) that contained either: pJU204 (WT, PRP8-3HA), pJU204ANLS, pJU204-rp5, pJU204-rp6, or pJU204-rp7 were grown in YPDA broth and analysed by immunoprecipitation. All mutated or deleted-NLS-Prp8p tested immunoprecipitated increased amounts of Aar2p when compared to the wild-type Prp8p (Figure 4.7A, lanes 6-10). Despite the fact that Brr2p levels were comparable, or increased in the input of tested mutated Prp8p extracts (Figure 4.7A, lanes 1-5), the pull-down of Brr2p by these mutated Prp8ps was obviously decreased (Figure 4.7A, lanes 6-10).

In addition, the association between mutated Prp8p and snRNAs were tested. An interesting finding is that mutated Prp8p (rp5) precipitated mainly one species of U5 L, and the U5S snRNA was present only in trace amounts (Figure 4.7B, lane 6). The reduced pull-
Figure 4.6. Prp8p C-terminus alignment, mutations, growth phenotype, and localisations. A) The sequences were identified using NCBI BLAST and then aligned using Clustal_W (1.82). The conserved amino acids are indicated with *; semi-conserved ones with (;) or (). Yeast Prp8p mutations (rp5:R2388K; rp6:R2388G, and rp7:F2392L) were created (according to human Prp8p mutations that resulted in Retinitis Pigmentosa type 13) by site-directed mutagenesis of the pJU204 plasmid that encodes full-length Prp8p. B) The WT, rp5, rp6, and rp7 plasmids were transformed into KL3-Brr2 yeast strain and plated on His dropout plates. The transformants were restreaked on YPDA plates and incubated at different temperatures (14°C, 23°C, 30°C, and 37°C). WT, and rp7 mutant grew at all tested temperature, but rp5 and rp6 were heat sensitive. C) The localisation of mutated Prp8p was investigated by indirect immunofluorescent staining. The nuclear localisation of Prp8p was not affected by the Prp8p C-terminus mutation and was identical to WT Prp8p (Figure 4.1A).
down of U4 and U6 snRNAs by mutated Prp8p (rp5) was most probably caused by weak association between Brr2p and Prp8p (Figure 4.7A, lane 8). The association between Brr2p and Prp8p to form the mature U5 snRNP (Stevens et al., 2001) is required prior to the formation of tri-snRNP. The observation that only U5-L snRNA is precipitated by mutated Prp8p (rp5) suggests that Prp8p can distinguish these two forms of U5 that are incorporated into the U5 snRNP. A similar finding was observed in another Prp8p C-terminal point mutation strain (prp8-I, G2347D). The prp8-I mutation not only disrupted the association between Prp8p and Brr2p (van Nues and Beggs, 2001), but the pull down of prp8-I protein showed no co-precipitation of U4, U5, or U6 snRNAs (Brown, 1992). The result of the ΔNLS-Prp8p snRNA precipitation was consistent with the previous co-IP assay (Figure 4.2C).

**Figure 4.7.** Immunoprecipitation of Prp8p revealed association between Prp8p and other U5 snRNP proteins and other snRNAs. A) The associations of Aar2p, Brr2p (13-Myc tagged at C-terminal), and Snul14p with Prp8p (WT and rp mutants) were analyzed by co-immunoprecipitation of HA-tagged Prp8p using anti-HA antibodies. The precipitates were fractionated on a 4-12% PAGE gel, blotted and probed with anti-Aar2p, anti-Myc, and anti-Snu114p antibodies. Deletion-NLS-Prp8p, and rp mutants showed increased Aar2p pull-down, but decreased Brr2p pull-down when compared to WT Prp8p pull-down. B) snRNAs co-precipitated by Prp8p were analyzed by Northern blotting. WT Prp8p precipitated U4, U5, and U6 snRNAs, but C-terminal mutated Prp8p (rp5) only precipitated one species of U5 (L) and trace amounts of U5S, U4, and U6. snRNAs pulled down by deletion-NLS-Prp8p are consistent with a previous observation (Figure 4.2C).
The distribution of Aar2p, Brp2p, and Snu114p in complexes was analysed, using a combination of glycerol gradient centrifugation and immunoprecipitation using Prp8p as bait. Alternate fractions from the glycerol gradient were immunoprecipitated with anti-HA agarose, and subjected to Western blotting. Co-precipitation of Aar2p was found to be increased in KL3-Brr2 strains that harboured pJU204ANLS or pJU204-rp5 (Figure 4.8B & 4.8C, fractions 1-13) when compared to the same strain that harboured pJU204 (WT) (Figure 4.8A, fractions 5-11).

Figure 4.8. Distribution of Prp8p associated complexes. Splicing extract (80 µl) was diluted with 120 µl GG buffer, and layered onto and sedimented through a 10-30% linear glycerol gradient and 400 µl fractions were collected. The odd-numbered fractions were incubated with 25 µl anti-HA agarose, on a rotating wheel for overnight, at 4°C. The precipitates were fractionated in a 4-12% PAGE gel, blotted and probed with anti-Aar2p, anti-Myc, and anti-Snu114p antibodies. WT Prp8p (A) showed increased levels of Snu114p that co-isolated with Brr2p at the bottom layer. Both Prp8p (rp5) mutant (B), and deletion-NLS-Prp8p (C) splicing extracts showed increased levels of Snu114p in the lower density fractions that co-isolated with Aar2p.
In the WT extract, Snu14p was more concentrated in the bottom fractions (Figure 4.8A, fractions 1&3) that co-localised with Brr2p, which is present only in the absence of Aar2p. In Prp8p-rp5, the distribution of Snu14p was different, being concentrated in fractions where Aar2p was present (Figure 4.8B, lane 7-11), and only trace amounts of Brr2p were associated with Prp8p (Figure 4.8B, fraction 1), even after a long film exposure. This could be due to the mutation of Prp8p disrupting the association of Brr2p, leading to accumulation of the Prp8p-Snu14p-Aar2p complex. Snu14p in ΔNLS Prp8p extract was evenly distributed between the fractions where Brr2p and Aar2p localised (Figure 4.8C, fractions 1-11). The even distribution of Snu14p could be due to Prp8p-Snu14p-Aar2p and Prp8p-Snu14p-Brr2p being in equilibrium. The defect in ΔNLSPrp8p-Snu14p-Aar2p import to the nucleus should accumulate substantial amounts of cytoplasmic ΔNLSPrp8p-Snu14p-Aar2p, however, unlike for Prp8p-rp5, the nuclear ΔNLSPrp8p-Snu14p-Aar2p complex does not have difficulty in Brr2p association.

The localisations of Brr2p and Prp8p in the prp8p-rp5 yeast strain were then investigated. Despite the previous report by Achsel et al. (1998), showing that Prp8p stably associated with Brr2p even in 0.2 M of the chaotropic salt sodium thiocyanate, immunoflourescent staining showed that Brr2p and Prp8p were not co-localised, even in the WT Prp8p strain (Figure 4.10A). Also, an increased level of cytoplasmic Prp8p signal was detected after prp8-rp5 was shifted for 2 h from 30°C to 37°C growth condition (Figure 4.9B and 4.9C). This observation may be similar to the report of increased level of point mutated PRPF31 in the cytoplasm, which could be due to the protein mis-folding, leading to insoluble aggregates in the cytosol (Deery et al., 2002). An interesting observation is that the prp8p-rp5 protein’s patchy staining towards the periphery of the nucleus was increased (Figure 4.10B) when compared to the WT Prp8p (Figure 4.9A).

The association and distribution of Brr2p was investigated by using KL1-Brr2 (Brr2p is tagged at the C-terminal with TAP, otherwise as KL1) harbouring either pJu204 (WT, PRP8-3HA), pJu204ΔNLS or pJu204-rp5, because Myc-tagged Brr2p (KL3-Brr2) was found poorly immunoprecipitated. The combination of glycerol gradient centrifugation and immunoprecipitation analysis was used as mentioned above. The precipitates were fractionated in 4-12% PAGE gel, blotted, and probed with anti-Snu114p, and anti-Aar2p antibodies. The Brr2p-TAP in the precipitates was detected due to Protein A binding IgG. Brr2p was found distributed in fractions 1-17 (Figure 4.10 A, B & C) regardless of the presence of WT or mutated Prp8p. A Brr2p that free from Snu114p was found in fractions 13-17, and Aar2p is constantly absent in all the Brr2p precipitates. Interestingly, Snu114p (Brr2p-free) was increased due to the Prp8p-rp5 mutation (Figure 4.10B, lanes 9-17), and an
increased portion of Brr2p was shifted to the lower density fractions (Figure 4.10B, lanes 11-17). On the other hand, ΔNLS-Prp8p caused Brr2p to be distributed more in fractions 5-11 when compared to the WT Prp8p. Yeast two hybrid results (van Nues and Beggs, 2001) showed that both N- and C-termini of Prp8p interact with Brr2p. The Brr2p relocated to lower density fractions could be due to the disruption of the Brr2p-Prp8p interaction caused by the Prp8p aa96-117 deletion, which happens to be the NLS of Prp8p.

**Figure 4.9. Prp8p and Brr2p localisations analysis.** Localisation of triple-HA tagged Prp8p and 13-Myc-tagged Brr2p was detected with rat anti-HA, and mouse anti-Myc primary antibodies respectively, samples were subsequently probed with Alexa fluor 594 anti-rat antibodies, and Alexa fluor 488 anti-mouse antibodies respectively. (A) Both WT Prp8p, and Brr2p showed sub-nuclear localisation, but were not co-localised to each other. (B) At 30°C growth conditions, the strain with mutant Prp8p (rp5) showed subnuclear localisation of Prp8p and Brr2p. C) When the rp5 strain was shifted from 30°C to 37°C for 2 h, Prp8p was detected in the cytoplasm, but Brr2p nuclear localisation was not affected by the shift of temperature.
Figure 4.10. Distribution of Brr2p associated complexes. Splicing extract (80 μl) was diluted with 120 μl GG buffer, and layered onto and sedimented through a 10-30% linear glycerol gradient and 400 μl fractions were collected. The odd-numbered fractions were incubated with 25 μl IgG-agarose, on a rotating wheel for overnight, at 4°C. The precipitates were fractionated in a 4-12% PAGE gel, blotted and probed with anti-Aar2p and anti-Snu114p antibodies. WT Prp8p (A) showed increased levels of Snu114p that co-isolated with Brr2p at the bottom layer. Both Prp8p- rp5 (B), and deletion-NLS-Prp8p (C) splicing extracts showed increased levels of Snu114p in the lower density fractions. Aar2p was not detected in the precipitates.
The results reported in the previous section suggest that U5 snRNP formation begins in the cytoplasm. This raised the question of how the cytoplasmic pre-U5 snRNP is imported into the nucleus. To investigate this, 3 importin-mutant yeast strains that show a defect in NLS protein nuclear import were tested. These are \textit{csel-1} (Solsbacher et al., 1998), \textit{kap95-ts/rsll-1} (Ueta et al., 2003), and \textit{NMD5} deletion strain (Boustany and Cyert, 2002).

The \textit{csel-1} (CS), \textit{kap95-ts/rsll-1} (TS), \textit{ANMD5} (no phenotype), and BMA38a (WT) yeast strains were first grown in YPDA broth at 30°C, and subsequently immunofluorescent stained with anti-Snu114p antibodies (Figure 4.11). BMA38a and \textit{csel-1} yeast strains showed defined nuclear staining of Snu114p (Figure 4.11A & 4.11B), while the trace amount of cytoplasmic signal detected could be possibly due to non-specific staining. In \textit{LINMD5} strains, a slight increase of cytoplasmic signal was detected when compared to BMA38a strain, but the nuclear signal is clearly detected (Figure 4.11D). In the a \textit{kap95-ts} strain, an increased proportion of Snu114p cytoplasmic signal was detected, and this cytoplasmic signal was observed in all the stained cells. Similar results were obtained by Nikolaev et al. (2003), who showed that Pdr3p was defective in nuclear localisation in \textit{kap95} mutant strain even the strain was grown at permissive temperature (Nikolaev et al., 2003).

I then further investigated the localisation of snRNAs (U2, U5, and U6), and a snoRNA (U3) by using \textit{kap95-ts} and WT strains grown at permissive (30°C) or non-permissive temperature (37°C). At both 30°C and 37°C, U2 is consistently localised in the nucleus of the WT strain (Figure 4.12A & 4.12B). In the \textit{kap95-ts} strain, a slight increase of U2 in the cytoplasm was detected when the cells were grown at 30°C (Figure 4.12C). And the U2 cytoplasmic signal was clearly increased when the \textit{kap95-ts} was shifted to 37°C for 2 h (Figure 4.12D).

Also, the U5 snRNA distribution was investigated. In a WT strain, U5 snRNA was detected predominantly in the nucleus and slight signal from the cytoplasm (Figure 4.13A & 4.13B). However, in the \textit{kap95-ts} strain, U5 snRNA signal was increased in the cytoplasm. This result is complicated by the fact that the distinction between nucleus and cytoplasm was not clear in the hybridised cells that were grown at either 30°C or 37°C growth conditions (Figure 4.13C & 4.13D). The snoRNA (U3) distribution was investigated in \textit{kap95-ts} strain. U3 nucleolar localisation was consistent in this temperature-sensitive yeast strain at either permissive or non-permissive temperatures (Figure 4.14A & 4.14B). Interestingly, nuclear localised U6 snRNA (Figure 4.14C) was delocalised to cytoplasm (Figure 4.14D) when
Snu114p distribution in importin-mutation yeast strains. Snu114p localisation was analysed in WT (BMA38a), cse1-1, kap95-ts, and ΔNMD5 strains. Yeast cells were probed with rabbit anti-Snu114p antibodies, and then with Alexa fluor 594 goat anti-rabbit secondary antibodies. A) WT yeast cells, and B) importin mutant, cse1-1 yeast cells showed nuclear localised Snu114p, with slight cytoplasmic signal that might be due to background staining. C) Yeast cells (kap95-ts) showed clear Snu114p signal from the cytoplasm. D) Yeast cells with NMD5 deleted showed a slight increase of Snu114p cytoplasmic signal, but may not be significant.
Figure 4.12. U2 localisation in nuclear import mutant *kap95-ts*.

Localisation of U2 snRNA in WT (W303a) and *kap95-ts* at 30°C and 37°C was detected with FISH. In the WT strain, U2 is consistently localised in the nucleus (A & B) in both 30°C and 37°C growth conditions. C) At 30°C, *kap95-ts* showed nuclear localised U2 snRNA, but also a detectable cytoplasmic U2 signal was observed. D) U2 was clearly detected in the cytoplasm when *kap95-ts* was shifted from 30°C to 37°C for 2 h.
Figure 4.13. U5 localisation in nuclear import mutant *kap95-ts*. Localisation of U5 snRNA in WT (W303a) and *kap95-ts* at 30°C and 37°C was detected with FISH. In the WT strain, U5 was detected in the predominantly in the nucleus, with a slight signal from the cytoplasm (A & B) at both 30°C and 37°C. At 30°C, (C) and 37°C (D) *kap95-ts* showed U5 snRNA increased cytoplasmic signal.
Figure 4.14. U3 and U6 localisations in nuclear import mutant *kap95-ts*. Localisation of U3 snoRNA in *kap95-ts* yeast strain was consistently nucleolar, regardless of growth at permissive (A) or non-permissive (B) temperature. In 30°C growth conditions, U6 showed nuclear localisation in *kap95-ts* yeast strain, but when this strain was shifted to 37°C for 2 h, the U6 was found in both nucleus and cytoplasm. E) The WT (W303a) U6 nuclear localisation was not affected by the shift of temperature.
In this chapter, I have identified that Prp8p aa96-117 encodes a functional NLS. Using a strain in which Prp8p is lacking this NLS, I looked into the localisation of other U5 snRNP proteins. Brr2p was consistently nuclear in both WT and ANLS Prp8p strains. In contrast, Snu114p is in the nucleus in the WT strain and was delocalised to the cytoplasm in the ANLS Prp8p strain. This suggests that Brr2p nuclear localisation is Prp8p-independent, but that Snu114p depends on Prp8p for its nuclear uptake or retention. The non-U5 snRNP components: U2 (Figure 4.4O & 4.4P), and U1, U6 snRNAs and the U6 core protein, Lsm8p, showed identical nuclear signals in both WT and ANLS Prp8p strains. In contrast, whereas U5 was predominantly in the nucleus in WT cells (Figure 4.3M), in the ANLS Prp8p strain, the U5 signal was increased in the cytoplasm (Figure 4.4N).

Previous publications showed that the U5 snRNP is present in at least two forms. By using Aar2p as bait, Gottschalk and co-workers isolated a core U5 snRNP (Gottschalk et al., 2001; Stevens et al., 2001). This core U5 snRNP consisted of Aar2p, Prp8p, Snu114p, Sm proteins, and U5 snRNA. At the same time, Stevens and co-workers isolated the more complex form of U5 snRNP by using Brr2p as bait. This U5 snRNP consisted of Brr2p, Prp8p, Snu114p, Snu40p, Prp28, Dib1p, Sm proteins, and U5 snRNA. As shown by Gottschalk et al. (2001), Aar2p functions as an U5 snRNP re-cycling factor, as depletion of Aar2p abolished the later round of splicing in vitro. Interestingly, Aar2p was not precipitated by biotinylated pre-mRNA, suggesting that this protein is not involved in the precursor or active spliceosome (Gottschalk et al., 2001). Also, the immunoprecipitation of Brr2p did not detect the presence of Aar2p. This suggested that Aar2p was released from the U5 precursor complex before Prp8p associated with Brr2p to form a more complex U5 snRNP. In this case, the prp8-rp5 mutation may have two effects on U5 precursor and mature snRNPs transition, in which the binding between Prp8p and Aar2p is increased, and at the same time, the binding between Prp8p and Brr2p is decreased. Due to the reports of these two U5 snRNP complexes, I decided to look into the localisation of Aar2p and the Sm ring members: SmB1p, SmD1p, and SmE1p. SmE1p was localised predominantly in the nucleus in the WT Prp8p strain, and showed a slight increase in cytoplasmic signal in ANLS Prp8p strain. However, the SmB1p and SmD1p, which contain NLSS (Bordonne, 2000) were consistently localised in the nucleus.

By immunoprecipitation of Prp8p, I showed that the levels of U4 and U6 snRNA that are bound to ANLS Prp8p were significantly down, unlike the level of co-precipitated U5, which was comparable in mutant and WT, despite the fact that ANLS-Prp8p was
predominantly cytoplasmic. The reduction of tri-snRNP in the ΔNLS Prp8p strain, when compared to WT Prp8p strain, was observed as well in the glycerol gradient analysis. This suggests that the U5 snRNA might associate with Prp8p-Snu114p to form a U5 precursor particle in the cytoplasm, that is imported into the nucleus where it forms a mature U5 snRNP complex with Brr2p. This supports the existence of a cytoplasmic phase during the biogenesis of U5 snRNA in budding yeast.

Aar2p was shown to be in both the nucleus and cytoplasm (Figure 4.4C). When Prp8p was over-expressed, Aar2p was found to be concentrated in the nucleus (Figure 4.5B, upper panel). When Prp8p was depleted, Aar2p was delocalised again to the cytoplasm (Figure 4.5B, lower panel). This suggests that Prp8p functions in nuclear retention, or mediates the nuclear import of Aar2p. In Chapter III, I showed that Aar2p association with Prp8p was disrupted by one of the partition points of Prp8p (aa2173). However, Aar2p was previously shown to be essential (Gottchalk et al., 2001), which suggests that the lethality of the yeast cell upon deletion of Aar2p may not be due to the disruption of association between Prp8p and Aar2p. As I proposed the formation of precursor U5 snRNP in the cytoplasm, I immunoprecipitated Aar2p and detected bound U5 snRNAs by Northern blotting. Despite the fact that ΔNLS Prp8p was delocalised to the cytoplasm, the U5 snRNA that was bound to Aar2p was increased in the pull-down when compared to the WT Prp8p pull-down (Figure 4.5A). The increased amount of U5 detected could be due to the increased amount of U5 precursor complex that is deficient in nuclear import due to ΔNLS Prp8p.

I also proposed that the mature U5 snRNPs are formed in the nucleus. Firstly, the nuclear localisation of Brr2p is not affected by the delocalisation of Prp8p (Figure 4.4I). I utilised the yeast Prp8p mutants (rp), which are identical to human Prp8p mutants that cause Retinitis Pigmentosa (RP13) to support this proposal. The nuclear localisations of these Prp8p mutants are not affected by the mutation when the cells are grown at 30°C (Figure 4.6C). This means that the precursor of Prp8p-Snu114p-Aar2p-U5 that forms in the cytoplasm has no defect in nuclear import. The results also showed that the rp mutants are reduced in Prp8p and Brr2p binding (Figure 4.7A, lanes 8-10). In this situation, the binding of Brr2p to Prp8p was reduced, and Aar2p was precipitated in an increased level by Prp8p (rp mutants), when compared to the pull-down by WT Prp8p (Figure 4.8B, fractions 5-11). The data presented suggest that the rp mutations cause a defect in mature U5 snRNP formation, and increased levels of precursor U5 snRNP that localise to the nucleus. That Prp8p and Brr2p did not co-localise was rather surprising (Figure 4.9A, B & C), and the distributions of Brr2p in glycerol gradient centrifugation was not affected much by prp8-rp5
protein (Figure 4.10B), suggesting that the association of Prp8p and Brr2p could be transient, despite the strong interaction that was shown previously (Achsel et al., 1998).

Subsequently, I identified Kap95p as the importin that is required for Snu114p and snRNA (U2, U5, and U6) nuclear import. The Kap95p human homologue, Importin β, is required for U snRNP import in metazoa (Palacios et al., 1997) and for proteins containing a classical NLS. Saturation of importin β-mediated transport with the importin β-binding domain of importin α blocks U snRNP import both in vitro and in vivo. Immunodepletion of importin β inhibits both NLS-mediated and U5 snRNP import (Palacios et al., 1997).

The following model is consistent with my observations (Figure 4.15). Prp8p-Snu114p-Aar2p-U5 complex exists as a U5 precursor particle in the cytoplasm, which is then imported into the nucleus, for which Kap95p is required. In the nucleus, Aar2p dissociates from the Aar2p-U5 complex, and subsequently the mature U5 snRNP complex with Brr2p is formed. The C-terminal region of Prp8p is required for this Brr2p association, as association between Prp8p and Brr2p is destabilised by the rp mutations in Prp8p.

Figure 4.15. Model for U5 snRNP assembly pathway. In this model, the formation of the U5 snRNP precursor complex that consists of Prp8p-Snu114p-Aar2p and U5 snRNA occurs in the cytoplasm. Deletion of the Prp8p NLS delays the nuclear import of this precursor U5 snRNP, thus immunoprecipitation and glycerol gradient analysis showed increased levels of this complex. The import of this complex is mediated by importin Kap95p. A temperature-sensitive mutant of kap95 showed increased cytoplasmic accumulation of Snu114p and U5 snRNA. Once the U5 precursor complex imported into the nucleus, Prp8p dissociates from Aar2p, and forms a more complex U5 snRNP with Brr2p. The C-terminal region of Prp8p is important in this association. Despite the fact that Prp8p C-terminus mutants (rp5, rp6, and rp7) showed no defect in Prp8p nuclear localisation, the mutants do show decreased Brr2p association and increased levels of Prp8p-Snu114p-Aar2p complex.
Chapter 5

Prp8p co-selected protein (Spp382p) analysis

5.1 Introduction

*SPP382* is an essential gene on chromosome XII of *S. cerevisiae* (Giaever *et al.*, 2002), and its product is present in the native splicing complexes (Pandit and Rymond, SGD). The 708 amino acid (aa) long Spp382p has a molecular weight of 83 kDa, and contains a G-patch domain at aa61-106 (Figure 5.1). The G-patch is found in a number of RNA binding proteins (Aravind and Koonin, 1999), which suggests that Spp382p may have an RNA binding function. Also, Spp382p was found on many occasions to be co-selected, or to have yeast-two-hybrid interaction with other splicing factors (Gavin *et al.*, 2002; Albers *et al.*, 2003; Hazbun *et al.*, 2003; Lebaron *et al.*, 2005). Indeed, TAP tag affinity selection of Prp8p partition fragment (aa771-2413) also co-selected this protein (Figure 3.4A; Boon *et al.*, in press in RNA). In addition, TAP tag purification using Spp382p as bait (Hazbun *et al.*, 2003) pulled down splicing factors from the U5 snRNP and the Prp19 complex.

By using NCBI blast, Spp382p was shown to be conserved throughout many eukaryotes; *Homo sapiens* (GI:21619831), *Canis familiaris* (GI:57105706), *Gallus gallus* (GI:53135754), *Mus musculus* (GI:10190660), *Rattus norvegicus* (GI: 56605648), *Xenopus laevis* (GI:51703472), and *Schizosaccharomyces pombe* (GI:6523773). The Spp382p homologues in mouse (TFIP11p) and in *S. pombe* (RAC; ribosomal assembly chaperone) show similarity with Spp382p throughout the protein (Figure 5.1). Recently, Wen *et al.* (2005) showed that mouse TFIP11 protein does not co-localise with any known splicing factors, indicating that the TFIP11 protein could have an as yet un-revealed function in splicing. Interestingly, the RAC complex exhibits no nuclease activity, and changes which disrupt RAC protein binding also inhibit Pac1 nuclease cleavage at the 3' end of the 25S rRNA sequence (Spasov *et al.*, 2002), which suggests a potential role for Spp382p in ribosomal RNA processing. The name *SPP382* was given to the gene *YLR424W* to indicate its function in suppression of the *prp38-1* mutation, which was found to be defective in spliceosome maturation (Xie *et al.*, 1998).

In this chapter, I demonstrate that Spp382p is spliceosome associated, as it co-precipitated U2, U5, U6, and excised intron. Depletion of Spp382p resulted in excised intron accumulation, and disruption of association of Prp43p and excised intron, which indicates that Spp382p is required for a late-stage of splicing.
Figure 5.1. The alignment of *S. cerevisiae* Spp382p with *M. musculus* (TFIP11p; GI:21619831) and *S. pombe* (RAC; GI:6523773) homologues. The sequences were identified using NCBI BLAST and then aligned using Clustal W (1.82). G-patch of *S. cerevisiae* Spp382p is underlined. According to Clustal W (1.82).
5.2 Co-immunoprecipitation of snRNAs and Spliceosome with Spp382p

The association of Spp382p to snRNA was tested with co-immunoprecipitation of the spliceosomal U1, U2, U4, U5, and U6 snRNAs with Spp382-TAP protein. Splicing extracts (50 µl) from strain KL4T (Spp382-TAP tag), and non-TAP (BMA38a: negative control) were incubated with IgG-agarose and the co-precipitated snRNAs were detected by Northern analysis (as mentioned in 2.3.7).

Under 150 mM salt precipitation conditions, Spp382-TAP tagged protein (KL4T) pulled down U2, U5 (L and S), and U6 snRNAs, but no U1 or U4 snRNAs (Figure 5.2A, Lane 4). As no ATP was added to the splicing extracts, this suggests that Spp382p associates with U2, U5, and U6 snRNAs under non-splicing conditions. The non-TAP extract that served as background control precipitated trace amounts of U1 (Figure 5.2A, lane 3), and the same U1 signal was observed in the tested sample.

In order to determine the association of Spp382p with the spliceosome and pre-mRNA, the Spp382-TAP tagged splicing extract was incubated under in vitro splicing conditions with radio-labelled pre-mRNA in a total of 60 µl splicing reactions (according to 2.3.5), and then the Spp382-TAP protein was precipitated with IgG-agarose. An aliquot of 10 µl splicing reaction mixture was analysed for splicing activity, while the remaining 50 µl of splicing reaction mixture was incubated with IgG-agarose. Splicing extract with Prp8p TAP-tagged was used as the positive control, and non-TAP splicing extract was used as the background control. Prp8p is a U5 snRNP protein associated with the spliceosome throughout the splicing reaction and should therefore precipitate the pre-mRNA, intermediates, and products of the splicing reaction (Figure 5.2B, lane 5), as shown previously by Teigelkamp et al. (1995).

Compared to the Prp8p control, Spp382-TAP tagged protein efficiently co-precipitated the pre-mRNA, and the excised intron (Figure 5.2C, lane 6). In extract without TAP-tagged protein, only very low levels of pre-mRNA was found in the precipitate (Figure 5.2B, lane 4), demonstrating the specificity of precipitation of Spp382p. These findings showed that Spp382p is associated with the spliceosome in the second catalytic step or later, when it is found associated with the excised intron. The precipitation of pre-mRNA by Spp382p could also suggest this protein could function as a recycling factor, which re-assembles the other splicing factors onto pre-mRNA before the reoccurrence of splicing.
Figure 5.2. Spp382p associate with snRNAs and intron RNA

(A) Co-precipitation of snRNAs with Spp382p. Splicing extracts (50 μl) with non-tagged (BMA38a; lane 3), or Spp382p-TAP (KL4T; lane 4) were mixed with an equal volume of precipitation buffer containing IgG-agarose beads and incubated at 4°C for 2 h. The beads were washed in buffer containing 150 mM NaCl, deproteinized, and the RNAs were precipitated, blotted and probed for U1, U2, U4, U5 and U6 snRNAs. Total RNA extracted from 1.25 μl of each extract is shown in lanes 1-3. (B) Co-precipitation of splicing complexes by Spp382p. Splicing extracts derived from non-tagged (BMA38a; lanes 1 and 4), Prp8p-TAP (RG8T; lanes 3 and 5), Spp382p-TAP (KL4T; lanes 3 and 6) strains were incubated under splicing conditions in vitro (60 μl total volume) with 32P-labeled ACT1 pre-mRNA as substrate. The reactions were stopped after 25 minutes and 10 μl were removed as the input controls. The remaining 50 μl were mixed with IgG-agarose beads and incubated at 4°C for 2 h. Beads were washed in buffer containing 150 mM NaCl, deproteinized, and the RNAs were immunoprecipitated (IP). The RNAs were analysed by PAGE and autoradiography. The various RNA species are indicated diagrammatically on the right of panel B, with rectangles representing exons, a thin line representing the intron, and a lariat/loop representing the branched form of the intron following the first catalytic step of splicing.
5.3 Depletion of Spp382p

KL4G was used for further functional characterisation of the Spp382 protein. An interesting finding is that, despite the fact that \textit{SPP382} is required for cell viability, KL4G that was grown under repressive conditions (YPDA) only caused slow growth and formation of heterogenous colonies (Figure 5.3A). To analyse the effect of \textit{SPP382} turned off by \textit{GAL} promoter, the KL4G was first grown in galactose medium and then shifted to glucose medium. As a control, a second culture was treated identically but was transferred back to galactose medium. Approximately two hours after the metabolic shift the growth of the culture was slowed considerably but the cells continued to grow slowly (approximately 4 hours doubling time).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.3.png}
\caption{A) KL4G was streaked on either a YPGR or YPDA plate, and incubated at 30°C for 4 days. KL4G grown on YPDA formed smaller, more heterogenous colonies when compared to KL4G that was grown on the YPGR plates. Growth rate of wild-type (BMA38a) and \textit{GAL::3HA-SPP382} (KL4G) strains following a transfer from permissive galactose medium to B) non-permissive glucose medium or C) galactose medium for the time indicated. Cells were grown at 30°C and maintained in exponential growth throughout the time course by addition of fresh medium.}
\end{figure}
However, the splicing extract derived from KL4G shifted from galactose medium to glucose medium for 10 hr showed efficient depletion of Spp382p by Western blot analysis (Figure 5.4B, lane 2) compared to KL4G that was grown in galactose medium (Figure 5.4B, lane 1). Nop1p protein that served as the loading control showed that approximately the same amount of depleted and non-depleted Spp382p splicing extract were loaded. In vitro splicing performed on Spp382p-depleted splicing extract showed accumulation of excised-intron, and reduced spliced mRNA (Figure 5.4C, Lane 6). Therefore, this suggests that Spp382p is required for excised intron metabolism in S. cerevisiae. Also, this suggests Spp382p could have a role in recycling the post-splicing components for a new round of splicing.

The distribution of snRNPs in extracts from Spp382p-depleted and non-depleted cells was analysed by fractionation in glycerol gradients followed by Northern blotting. The most-dense fractions contained a mix of U4/U6.U5 tri-snRNPs, spliceosomes and/or post-splicing complexes (Fig. 5.5A and B, fractions 1 to 5), then free U5 snRNP (fractions 5 to 9), U4/U6 di-snRNPs (fractions 9-13), and free U6 snRNP (fractions 15-19).

Figure 5.4. Effects of Spp382p depletion on splicing in vitro. KL4G cells were grown in YPGR medium to OD$_{600}$ 0.5, then spun down, washed with YPDA, respun, and resuspended in the original culture volume of YPDA and grown for 10 hours (the cells were maintained at log phase by adding fresh YPDA) before harvesting to produce a splicing extract. A control culture of KL4G was treated identically but was returned to galactose medium instead of glucose. A) Western blot showing the relative amounts of 3HA-tagged-Spp382p in KL4G cells at 0 h (lane 1) and at 10 hours (lane 2) after the shift to YPDA. Nop1p was probed as a loading control. B) In vitro splicing assays with extracts derived from Spp382p-depleted (lanes 4 to 6) and non-depleted (lanes 1 to 3) cells. Aliquots (5 µl) were withdrawn and halted after the incubation for 0, 10, and 25 min at 23°C. The RNAs were analyzed by PAGE and autoradiography.
In Spp382p-depleted extract there was a clear increase in the amount of U4 snRNA present in U4/U6 di-snRNP (Fig. 5.5B, fractions 9-13), whereas the levels of free U5 snRNP (fractions 5 to 9), and free U6 snRNP (fractions 15-19) were reduced. The reduction of free U5 and U6 snRNPs (and presumably of free U2 snRNP, although this cannot be assessed with these gradients) and the accumulation of U4/U6 di-snRNPs suggests that free U5 and U6 snRNPs were not regenerated, as a consequence of the defective disassembly of post-splicing U2/U5/U6 snRNP complex. Thus free U5 snRNP became the limiting factor for the conversion of U4/U6 di-snRNP to U4/U6.U5 tri-snRNP (as U6 snRNP is in excess), and presumably tri-snRNP and/or free U2 snRNP became limiting for spliceosome assembly.

Figure 5.5. Depletion of Spp382p leads to reduced free U5 and U6 snRNPs, but increased U4/U6 di-snRNPs and U2/U5/U6 snRNP complex. To examine the profile of snRNP distribution, (A) Spp382p non-depleted, and (B) Spp382-depleted extracts were centrifuged through 10–30% glycerol gradients and the snRNA contents of alternate fractions were analyzed by Northern blotting. Blots of the extracted RNAs were hybridized with probes for U1, U2, U4, U5, and U6 snRNAs. The sedimentation of U5, U6, and U4/U6 is indicated. *Fractions 1 to 7 contain a mix of large complexes, including spliceosomes, U4/U6.U5 tri-snRNPs and post-splicing complex. C) Co-precipitation of snRNAs by Prp8p in extracts from Spp382p-depleted (lanes 4 to 6) or non-depleted (lanes 1 to 3) cells. Pre-immune (PI) serum was used as background control.
Next, the Spp382-depleted and non-depleted extracts were incubated with anti-Prp8p antibodies. As expected, anti-Prp8p antibodies efficiently pulled down U5 snRNP and U4/U6.U5 tri-snRNP from non-depleted extract (Figure 5.5C, lane 3). With Spp382-depleted extract, immunoprecipitation of Prp8p pulled down an eight-fold increased amount of U2 (compared to non-depleted extract and normalized against the total amounts in each extract), but only half the amount of U4 (Figure 5.5C, lane 6), and an equivalent amount of U6. The increased association of Prp8p with a U2 snRNA-containing complex and decreased association with U4 snRNP in Spp382-depleted extract most likely reflects accumulation of Prp8p in the U2/U5/U6 post-splicing excised intron-containing complex.

5.4 Spp382p distribution and association with Prp43 protein

The distribution of Spp382p in splicing extract derived from KL4G (strain with SPP382 expressed under Pgal1 promoter and triple-HA fused at the 5' end) and KL4T (strain with SPP382 TAP tag at the C-terminal) was investigated by glycerol gradient centrifugation (as described in 2.8). The glycerol gradient fractions were immunoprecipitated by anti-HA (KL4G) or IgG (KL4T) agarose, fractionated in a 4-12% PAGE gel, blotted, and probed with anti-HA (KL4G), or PAP for KL4T. Spp382p was found distributed in a broad range of fractions (Figure 5.6 A&B, lanes 1-21). The extra bands (Figure 5.2A, lanes 5-21) below Spp382p (*) were also observed in KL4T (strain with SPP382-TAP tagged) derived extracts (Figure 5.6 B, lanes 5-21). This could be an isoform of Spp382p, as the Spp382p mouse homolog (TFIP11p) was highly phosphorylated (Wen et al., 2005).

Prp8p and the excised intron release factor, Prp43p, are amongst many proteins that have been found in affinity-selected complexes along with Spp382p (Gavin et al., 2002; Hazbun et al., 2003). To investigate this further, extract from KL4G cells was fractionated on a glycerol gradient and alternate fractions were analysed by precipitation with anti-HA antibodies, which showed co-precipitation of Prp8p and of Prp43p with 3HA-Spp382p in the same high density fractions 1 to 9 (Fig. 5.6C). As Spp382p pulls down Prp8p and Prp43p in the same high density fractions, it is most likely that these 3 proteins are present in the same high molecular weight complex. In addition, it is notable that the bulk of Spp382p and Prp43p occur in lower density fractions (Figure 5.6C, fractions 13-21). This likely represents an Spp382p/Prp43p complex that is non-snRNP-associated. Indeed, an interaction between Spp382p and Prp43pT123A could be demonstrated in the absence of other yeast proteins (Figure 5.6D). The Prp43pT123A was an ATPase-defective Prp43p that interfered in trans with
the *in vitro* splicing function of wild-type Prp43. Prp43<sub>T123A</sub> did not affect the chemical steps of splicing or the release of mature mRNA from the spliceosome. However, it blocked the release of the excised lariat-intron from the spliceosome (Martin *et al.*, 2002).

![Glycerol gradient fractionation of Spp382p](image)

**Figure 5.6.** Glycerol gradient fractionation of Spp382p. KL4G cell extract containing **A)** triple-HA tagged Spp382p or **B)** C-terminal TAP tag was fractionated in a 10-30% glycerol gradient. Alternate gradient fractions were immunoprecipitated with (A) anti-HA agarose or (B) IgG agarose. The precipitated proteins were analyzed by Western blotting with (A) anti-HA or (B) PAP. Asterisk (*) indicates the presence of alternate Spp382p bands. **C)** KL4G cell extract was fractionated in a 10-30% glycerol gradient and immunoprecipitated with anti-HA agarose as mentioned in (A), and the precipitated proteins were analyzed by Western blotting with affinity purified anti-Prp43 antibodies and with anti-Prp8p antibodies. **D)** Prp43<sub>T123A</sub> protein was incubated with recombinant GST-Spp382p or with GST alone at a concentration of 40 μg of each protein/ml. Following incubation with glutathione-Sepharose in IPP buffer (6 mM HEPES, pH7.9, 150 mM NaCl, 5 mM MgCl₂, 0.1% (v/v) Nonidet P40), the beads were washed, incubated with SDS loading buffer. The precipitates were fractionated on a 4-12% PAGE gel and Western blotted, and probed with affinity purified rabbit anti-Prp43p antibodies, and rabbit anti-GST antibodies. (The Prp43<sub>T123A</sub> protein is a stock protein in Beggs lab that was expressed and purified by Mr. A. Droop during his honours project in year 2003, Beggs lab).
5.5 Spp382p mediates Prp43p function in post-splicing complex

Figure 5.7 is provided by Dr. Gretchen Edwalds-Gilbert (a collaborator on the study of Spp382p).

Martin et al. (2002) produced an ATPase defective mutation T123A in conserved motif I of the Prp43 RNA helicase that confers a dominant negative phenotype. The Prp43_{T123A} protein interfered in trans with the in vitro splicing function of the wild-type Prp43 protein, blocking release of the excised-intron from the spliceosome (Martin et al., 2002). We used purified, recombinant Prp43_{T123A} protein to test whether Prp43p is able to associate with the post-splicing excised intron complex produced in reactions with Spp382-depleted and non-depleted extracts. As a control, anti-Prp8 antibodies were used to test for effects of Spp382-depletion on the association of Prp8p with the excised intron. In Spp382-depleted extract, with or without addition of recombinant Prp43_{T123A} protein, Prp8p co-precipitated excised-intron (Figure 5.7, lane 1-4). In contrast, Prp43_{T123A} protein only co-precipitated excised-intron from cell extract containing Spp382p (Fig. 5.7, lane 5).

These results show that the presence of Spp382p is required for the interaction of Prp43p with the post-splicing excised intron complex and explains the observed accumulation of excised intron complex in Spp382-depleted extract.

<table>
<thead>
<tr>
<th></th>
<th>Co-IP by anti-8.6 antibodies</th>
<th>Co-IP by anti-Prp43p antibodies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spp382p</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Prp43p_{T123A}</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

Figure 5.7. Splicing extracts with depleted or non-depleted 3HA-Spp382p were incubated under splicing conditions in vitro with ^32P-labeled ACT1 pre-mRNA as substrate. The reactions were stopped and incubated at 4°C for 2 h with anti-8.6 or anti-Prp43p (antibodies) absorbed sepharose A beads in 150 mM NaCl. The precipitates were deproteinized, and the RNAs were analysed by PAGE and autoradiography.
Using endogenously expressed or over-expressed Spp382p yeast strains, the localisation of Spp382p in relation to its expression level was tested. The endogenous Spp382p-expressing strain was constructed by tagging the C-terminal of Spp382p with the TAP tag. By indirect immunoflorescent, Spp382p was shown to be nuclear localised (Figure 5.8A): the 496-diamidino-2-phenylindole (DAPI) staining of DNA merged with the immunostaining obtained with anti-Protein-A primary antibodies. The over-expressing Spp382p strain (YLR424W::nat, [pGFP-N-424], otherwise as BMA38a) was constructed by inserting a plasmid (pGFP-N-424) encoding SPP382 with a GFP tag at its N-terminus, and under the control of P\textsubscript{MET23} promoter, while the endogenous SPP382 was deleted with natMX6. Surprisingly, the localisation of Spp382p was sub-nuclear (Figure 5.8B).

The mouse homolog of SPP382, TFIP11, was recently shown to be sub-nuclear localised in transfected cells (HEK293) (Wen et al., 2005). However, TFIP11 protein showed no co-localisation with any splicing factors tested so far: SC35 nuclear speckles, paraspeckles, Cajal bodies, Sm, hnRNP, PML, and Sam68 (Wen et al., 2005).

The in vitro co-precipitation (Figure 5.6D) showed that Spp382p interacts directly with Prp43p in the absence of other yeast proteins. To further the localisation analysis, I decided to investigate the localisation of Prp43p in relation to Spp382p expression level. Normally, Prp43p is localised in the nucleolus (SGD). Interestingly, in Spp382p overexpressed cells (KL4G), Prp43p showed aggregated sub-nuclear or sub-nucleolar staining (Figure 5.8C). When Spp382p expression is repressed (KL4G galactose to glucose shift for 4 hr), the Prp43p was relocalised to the nucleolus (Figure 5.8D).

The Prp43p foci could potentially be part of a complex containing Spp382p, as overexpressed GFP-Spp382p also showed aggregated nuclear staining. In addition, the aggregates could also be present in the nucleolus (Figure 5.8C), but further localisation with a nucleolar marker is needed for this hypothesis. Interestingly, depletion of Spp382p in strain KL4G by galactose to glucose medium shift (2 hr) showed accumulation of 35S rRNA, whereas the parent strain (BMA38a) did not show any 35S rRNA accumulation (Christophe Dez, unpublished data). This suggests that Spp382p could potentially function in both rRNA processing and spliceosome processing. However, the accumulation of 35S rRNA in Spp382p-depleted cells could be due to the secondary defect of defective of splicing.
Figure 5.8. Spp382p localisations. A) Endogenous Spp382p-TAP expressing yeast cells (KL4T) immunofluorescent staining revealed Spp382p nuclear localisation. B) Over-expressed GFP-Spp382p yeast cells showed GFP-Spp382p sub-nuclear localisation. C) Prp43p was probed with affinity purified rabbit anti-prp43p antibodies, and then detected by anti-rabbit Alexa-fluor594 secondary antibodies. Prp43p showed aggregate nuclear staining in 3HA-Spp382p overexpressed yeast cells (KL4G that grewed in YPGR). D) Prp43p was detected as mentioned in C), Prp43p was showed nucleolar staining in yeast cells (KL4G) that were shifted from YPGR to YPDA for 4 hr.
5.7 Discussion

The precipitation of U2, U5, U6 (Figure 5.2A), and mainly excised intron (Figure 5.2B) by Spp382p suggests that the protein is present in post-splicing complex. Indeed, the depletion of Spp382p not only caused accumulation of excised intron following splicing in vitro (Figure 5.4B, lane 6), but also reduced the levels of free U5 and U6 snRNPs (Figure 5.5B, lanes 5-7 & lanes 15-19). In contrast, the levels of U4/U6 di-snRNPs were increased (Figure 5.5B, lanes 9-13). The snRNAs in splicing extracts derived from Spp382p-depleted cells showed increased levels of U2, but decreased levels of U4, pulled down by Prp8p (Figure 5.5C, lane 6). All these data suggest that Spp382p is needed for snRNP regeneration.

By micro-array analysis (Barrass and Beggs, unpublished data), Spp382p-depleted cells showed an increase of unspliced pre-mRNAs. Using a new method of micro-array analysis that targeted the lariat-intron structure of mRNA, David Barrass (unpublished results) showed that the lariat-intron forms of mRNA were increased in vivo when Spp382p was depleted (the depletion of Spp382p was performed as mentioned in section 5.5).

Spp382p-TAP purification showed that this protein associated with U5 snRNPs and the Prp19 complex (Hazbun et al., 2003). Indeed, this protein co-purified with Prp8p (Figure 5.6C, lanes 1-9). In addition, this protein also interacts with Prp43p (Hazbun et al., 2003; Gavin et al., 2002), which was previously reported as a post-splicing complex disassembly factor (Arenas and Abelson, 1997; Martin et al., 2002). Interestingly, McPheeters and Muhlenkamp (2003) reported an unidentified ~80-kDa protein cross-linked near the branch site sequence upon completion of the second step of splicing. The predicted size of Spp382p is 83 kDa, and its function in late stage splicing and excised intron precipitation suggests the potential for Spp382p to be the unidentified protein, but this would need more biochemical proof in the future.

Yeast cells with the \textit{SPP382} gene under the control of regulatable \textit{P\textsubscript{GALJ}} promoter were shown to be able to grow slowly in glucose medium (Figure 5.3A). A similar situation was found in cells with the \textit{PRP43} gene under the \textit{P\textsubscript{GAL}} promoter, in which Prp43p was not efficiently depleted under the galactose-regulated promoter (Martin et al., 2002). This suggests that yeast cells could survive with only minute amounts of Spp382p, possibly due to leaking of the \textit{P\textsubscript{GALJ}} promoter.

Spp382p was named after the role of this protein in \textit{prp38-1} suppression (SGD). Prp38p is the first tri-snRNP-specific protein that has been shown to be dispensable in spliceosome assembly, but required for spliceosome maturation (Xie et al., 1998). The depletion of Prp38p caused a splicing complex maturation defect with accumulation of
U4/U6, and the temperature sensitive prp38-1 strain accumulated RP51A pre-mRNA in vivo at non-permissive temperature (Xie et al., 1998). The suppression of prp38-1 by Spp382p suggests an additional role(s) of this protein in splicing complex maturation.

Spp382p and Prp43p co-fractionated in snRNP-free fractions from a glycerol gradient, therefore they can apparently co-exist in a protein complex that may or may not contain other proteins. As the purified recombinant GST-Spp382p interacts with purified recombinant Prp43p<sub>7123A</sub>, this is a direct interaction that does not require any other yeast protein. Thus Spp382p is a candidate cofactor for the Prp43 helicase and apparently targets it to the post-spliceosome complex. Spp382p contains a G-patch domain (Aravind and Koonin, 1999) near its N-terminus (amino acids 61 to 106) that is the most highly conserved region of this protein. It will be interesting to determine whether the G-patch in Spp382p binds RNA and functions to promote the interaction of Prp43p with the intron RNA in the post-spliceosome complex.

As a splicing factor, Spp382p nuclear localisation (Figure 5.8A) is expected. But the detection of Spp382p in sub-nuclear bodies (Figure 5.8B) when this protein is over-expressed is rather surprising. Furthermore, these sub-nuclear bodies did not co-localise with any other known splicing bodies in yeast or mammalian cells (Wen et al., 2005) indicates this could be a new uncharacterized splicing structure. However, yeast two hybrid results (Tsai et al., 2005) showing that Spp382p bait interacted with Spp382p prey can give a different explanation. This suggests that Spp382p could self-interact, and the sub-nuclear bodies could be due to the self-interacting Spp382p aggregates. The aggregated staining of Prp43p in Spp382p overexpressed condition (Figure 5.8C), and the fact that these two proteins could directly interact without any additional yeast protein suggests that the aggregates of Prp43p could be in a complex with Spp382p. However, further co-localisation study is needed to confirm this assumption.
Chapter 6
Final Discussion & Future Work

The Prp8 protein is at the catalytic centre of the splicing machinery (section 1.7). Despite the fact that identification and analysis of this protein began in the 1980s, the complete function and mechanism of this protein in modulation of pre-mRNA splicing still remain to be elucidated. The large size of this protein with its lack of obvious domains, and its conserved homology across eukaryotes, emphasised the importance of analysis and understanding of this protein.

In this study, I have utilised Prp8p partitioned clones (chapter 3) to reveal the protein-protein interaction sites between Prp8p and related splicing factors. This led me to identify that the Prp8p 437-770aa polypeptide associates with Snu114p. Also, partitioning of Prp8p at position 2173 amino acid disrupts the Prp8p-Aar2p association. In addition, I have confirmed that the association of Snu114p and Aar2p with Prp8p are not interdependent (Boon et al., in press in RNA).

Assembly of the metazoan U snRNPs involves a multi-step process (section 1.6). On the other hand, the yeast U snRNP assembly pathway is still unclear. In chapter 4, I proposed a model for the yeast U5 snRNP assembly pathway: Prp8p associates with Snu114p, Aar2p, and U5 snRNA in the cytoplasm, the import of this cytoplasmic U5 snRNP is mediated by Kap95p, and the association of Prp8p and Brr2p occurs in the nucleus. In addition, Prp8p C-terminus mutations that resemble human Prp8p mutations in Retinitis Pigmentosa (RP) type 13 decreased the association between Prp8p and Brr2p, and the RP mutations in yeast Prp8p resulted in increased levels of the Prp8p-Snu114p-Aar2p complex.

The affinity purification of Prp8p co-selected an uncharacterised protein, Spp382p (chapter 3, figure 3.4A). In chapter 5, I showed that the Spp382 protein immunoprecipitated the U2, U5, U6 snRNAs, and excised-intron. Furthermore, the metabolic depletion of this protein resulted in accumulation of excised-intron. The Spp382 protein showed direct interaction with Prp43p, which is also a protein involved in late-stage of splicing (Martin et al., 2002). Interestingly, excised intron was not associated with Prp43p in the absence of Spp382p, but remained in a Prp8p-associated complex (Edwalds-Gilbert and Beggs, unpublished data). This suggested that Spp382p could function as a receptor, or by modulating the position of Prp43p in the late-stage of splicing to target the excised intron upon completion of the second step of splicing.
6.1 Future work – Part one

A U5 snRNP assembly pathway model was proposed based on the analyses performed in chapter 4 (Figure 4.15). Interestingly, Prp8p co-selected Mlp2p through TAP tag purification (KL-Boon and D. Short, unpublished data). And both Prp8p and Mlp2p showed punctate perinuclear staining (Galy et al., 2004). Increased levels of Prp8p-Snu114p-Aar2p complex accumulate in the nucleus in prp8-rp5 cells, and this complex might co-localise with Mlp2p in the peri-nuclear region. If true, this would suggest that Mlp2p might serve as a platform for Prp8p (Prp8p-Snu114p-Aar2p) before the formation of mature U5 snRNP with Brr2p.

Interestingly, Mlp2p was shown to have the function of pre-mRNA nuclear retention (Galy et al., 2004). However, Mlp2p was not known to have any RNA binding function. On the other hand, Prp8p was previously shown crosslinked to pre-mRNA 5' splice site even in the absence of U5 loop 1 (which anchors the exons of pre-mRNA in splicing in vitro; section 1.7). Based on the information stated above, Mlp2p could potentially dock Prp8p and its associated complex at the edge of the nucleus and, later Prp8p could prevent the leakage of pre-mRNA from the nucleus by binding pre-mRNA that is close to the pore site and subsequently proceed for pre-mRNA splicing. A potential model is proposed in the figure below.

![Figure 6.1.](image)

The proposed model for future analysis of the U5 snRNP assembly pathway.
6.2 Future work – Part two

In chapter 5, I analysed the Prp8 protein co-selected protein, Spp382p. Later, this protein was shown to function in the late-stage of splicing. Interestingly, in Spp382p-depleted extract, Prp43p was not able to pull down the excised-intron (Edwalds-Gilbert and Beggs, unpublished data). Despite the fact that Prp43p is needed for post-splicing complex disassembly through its helicase activity (Martin et al., 2002; Arenas and Abelson, 1997), the direct binding of Prp43p to RNA substrate (snRNA, pre-mRNA, or excised intron) was not shown. Interestingly, Spp382p includes a G-patch domain (aa61-106), and showed a direct binding to Prp43p in the absence of other yeast proteins (Figure 5.6D). Therefore, it is possible that Spp382p binds to excised-intron, or even base-paired U2/U6 in the post-splicing complex, and then serves as a platform for Prp43p, so that this protein can target the excised-intron or U2/U6 for unwinding or disassembly. The model for this hypothesis is proposed in the figure below.

![Figure 6.2.](image)

The propose model of future work for analysis of the function of Spp382p in orienting Prp43p in U2/U6 or excised-intron release.
References


Seraphin, B. (1995). Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. EMBO J 14, 2089-2098.


van Nues, R. W., and Beggs, J. D. (2001). Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of Saccharomyces cerevisiae. Genetics 157, 1451-1467.


Prp8p dissection reveals domain structure and protein interaction sites

KUM-LOONG BOON,¹ CHRISTINE M. NORMAN,² RICHARD J. GRAINGER,¹ ANDREW J. NEWMAN,² and JEAN D. BEGGS¹

¹Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom
²MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, United Kingdom
Prp8p dissection reveals domain structure and protein interaction sites

KUM-LOONG BOON,1 CHRISTINE M. NORMAN,2 RICHARD J. GRAINGER,1 ANDREW J. NEWMAN,2 and JEAN D. BEGGS1

1Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, United Kingdom
2MRC Laboratory of Molecular Biology, Cambridge CB2 2QH, United Kingdom

ABSTRACT

We describe a novel approach to characterize the functional domains of a protein in vivo. This involves the use of a custom-built Tn5-based transposon that causes the expression of a target gene as two contiguous polypeptides. When used as a genetic screen to dissect the budding yeast PRP8 gene, this showed that Prp8 protein could be dissected into three distinct pairs of functional polypeptides. Thus, four functional domains can be defined in the 2413-residue Prp8 protein, with boundaries in the regions of amino acids 394-443, 770, and 2170-2179. The central region of the protein was resistant to dissection by this approach, suggesting that it represents one large functional unit. The dissected constructs allowed investigation of factors that associate strongly with the N- or the C-terminal Prp8 protein fragments. Thus, the 1J5 snRNP protein Snu114p associates with Prp8p in the region 437-770, whereas fragmenting Prp8p at residue 2173 destabilizes its association with Aar2p.

Keywords: protein domains; protein interaction; RNA splicing; snRNP; transposon mutagenesis

INTRODUCTION

Studies of the structure, function, and molecular interactions of very large proteins are often hindered by the difficulty in obtaining sufficient quantities of the intact protein. On the other hand, many proteins appear to have been assembled from several independently folding domains, each of which has a distinct function, such as catalysis, ligand-binding, protein or nucleic acid interaction, or membrane-spanning properties (Doolittle 1995). Thus, the ability to identify and isolate individual protein domains offers a reductionist approach to facilitate protein characterization. However, in the absence of clear structural features it may be difficult to recognize domain structure in a protein. We describe here a novel approach to characterize the functional domains of a protein in vivo, using a custom-built Tn5-based transposon that causes the expression of a target gene as two contiguous polypeptides that must function in trans. We used this approach to investigate the domain structure of the large (280 kDa in budding yeast) spliceosomal protein, Prp8p. Prp8p is known to have multiple interactions in ribonucleoprotein (RNP) complexes, including the spliceosome (for review, see Grainger and Beggs 2005); however, despite the high level of conservation of this essential splicing factor, informatic analysis has provided few clues as to its likely domain structure.

The spliceosome is a dynamic complex that catalyses nuclear pre-mRNA splicing. It assembles on the pre-mRNA in a series of steps, in which the U1, U2, U4, U5, and U6 snRNAs in the form of small RNP particles (snRNPs) and a variety of non-snRNP proteins associate with the pre-mRNA. Prp8p is the largest and most highly conserved spliceosomal protein. Biochemical, genetic, and two-hybrid interaction data indicate that Prp8p interacts physically and/or functionally with a large number of splicing factors. It is a component of the U5 snRNP, a U5 snRNP/Prp19p complex, the U5-U4/U6 tri-snRNP, and the spliceosome (Grainger and Beggs 2005). It contacts the catalytic center of the spliceosome, as detected by photochemical cross-linking to functionally important regions of the U5 and U6 snRNAs and to the 5' splice site, the 3' splice site, and the branch point of the pre-mRNA (for review, see Grainger and Beggs 2005). In addition,
Prp8p has been implicated as a regulator of several conformational rearrangements in the spliceosome (Newman 1997; Staley and Guthrie 1998; Collins and Guthrie 1999; Brow 2002; Kuhn et al. 2002). Achsel et al. (1998) showed that human Prp8p (hPrp8p) forms a stable RNA-free core complex with h116K (ySnu114p), h200K (yBrr2p), and h40K (no known yeast homolog). Thus Prp8p may function as a large protein scaffold in the spliceosome and play a key role at its catalytic center.

Further characterization of Prp8 protein function(s) would be facilitated by knowledge of the domain structure of this very large 2413-residue protein. However, the sequence of Prp8p is largely novel and reveals little about the protein’s domain structure or its possible biochemical activities. We show here, using a novel random transposon insertion strategy, that the PRP8 gene of Saccharomyces cerevisiae can be dissected at any one of three positions (amino acids 394–443, 770, and 2170–2179), such that the yeast Prp8p can function as pairs of separate, interacting polypeptides. These dissected constructs allowed us to identify amino acids 437–770 of Prp8p as being associated with Snu114p, whereas disruption of Prp8p at 2173 interferes with its association with Aar2p. The central region of Prp8p, from amino acids 770 to 2173, was resistant to dissection by this approach, suggesting that it represents one large functional unit.

RESULTS

Transposon insertion to dissect PRP8

Our approach involves the random insertion of a custom-built Tn5-based transposon into the PRP8 gene to create a library of thousands of prp8 genes, each with a transposon sequence interrupting the open reading frame. The transposon sequence is designed to terminate translation of the upstream reading frame and at the same time drive expression of the prp8 sequence downstream of the site of insertion. The transposon consists of four elements (Fig. 1) inserted between the “Mosaic End” recognition sequences for Tn5 Transposase: (1) PRP8 gene downstream flanking region; (2) Kan or KanMX4 cassette to mark the transposon in bacteria or yeast, respectively; (3) PRP8 gene upstream flanking region, including the promoter; and (4) translation start site upstream of an SV40 nuclear localization signal (NLS). Insertion of this transposon will generally be lethal for five out of the six possible reading frames because in functional terms the gene will be truncated. However, in the remaining frame the transposon will direct the synthesis of an NLS-tagged downstream polypeptide fragment, resulting in two separate polypeptides encoded by distinct mRNAs.

We reasoned that if Prp8p does indeed have a modular structure, a small fraction of transposon-dissected genes might be functional. For example, if the transposon inserted in a region between essential functional modules, the two Prp8p fragments might fold independently and associate with each other and/or with other spliceosome components to perform the protein’s essential functions. We screened for functional transposon-dissected alleles of prp8 via plasmid shuffle in a yeast strain carrying a deletion of the chromosomal copy of PRP8. An estimated 60,000 transposition events occurred, of which at least 40,000 were located within the PRP8 open reading frame (i.e., an average of five
to six transposition events per nucleotide of coding sequence), and DNA digests showed a broad spread of insertion points (data not shown). As expected, the majority of prp8 alleles carrying a transposon were unable to support growth, but a small number did survive the selection. The prp8::Tn plasmids from these strains were isolated and restested, and the transposon locations were determined by restriction mapping and DNA sequencing.

Several transposons mapped within the first 80 residues of Prp8p, a region containing multiple Proline-rich tracts that are found only in plant and fungal Prp8 proteins (Grainger and Beggs 2005) and are not essential in yeast (K.-L. Boon and J.D. Beggs, unpubl.); these alleles were not analyzed further. Of the remaining gene dissection points, several mapped close together in the vicinity of amino acids 394–443 (specifically, after codons 394, 413, 416, 428, 436, and 443), and there were functional alleles with transposons inserted after codons 770, 2170, and 2179. Significantly, we screened through many thousands of transposons in the interval between 770 and 2170 in Prp8p in an effort to identify additional viable alleles, but none emerged. Western blotting confirmed that Prp8p fragments of the expected sizes were produced in place of full-length Prp8p, and sequence analysis showed that the transposon was inserted in the correct frame to drive expression of an NLS-tagged downstream Prp8p fragment (data not shown).

Finally we found by subcloning fragments of dissected pip8 genes to separate plasmids and transforming them individually into yeast cells that for each dissected prp8 allele, expression of both the N- and the C-terminal Prp8p fragments was required for growth (data not shown).

**TAP-tag purification of Prp8p complexes**

Yeast strains were constructed in which PRP8 is C-terminally TAP-tagged (Prp8T) and dissected immediately after the TAP affinity-selection procedure (Rigaup et al. 1999) and the coselected proteins were separated by SDS-PAGE and analyzed by mass spectrometry. Proteins that were coselected with full-length Prp8p included Aar2p, Brr2p, Prp3p, Prp4p, Prp6p, and Snu14p, and several Sm and Lsm proteins (Fig. 2A). These represent a subset of the proteins that were previously shown to be components of the U5 snRNP or U4/U6.U5 tri-snRNP (Gottschalk et al. 1999; Stevens et al. 2001, 2002). The Prp8T/770p C-terminal fragment (771–2143) coselected Aar2p, Brr2p, Prp3p, Prp4p, Prp6p, and Spp382p. The Prp8T/2173p C-terminal (2174–2413) coselected a protein similar in size to Brr2p, whereas other coselected proteins were in very low abundance due to the reduced stability of this complex. Association of Brr2p with the extreme C terminus of Prp8p was observed previously in two-hybrid interaction studies (van Nues and Beggs 2001).

Interestingly, Snu14p was hardly detectable in association with Prp8T/770p, whereas the amount of Aar2p associated
associated with Prp8T/770p was increased compared to the full-length Prp8p. Aar2p was previously found associated with a 16S form of yeast U5 snRNP that contained Prp8p, Snu114p, and the Sm proteins but lacked other known Prp8-associated proteins, including Brr2p. It was suggested that the Aar2-U5 snRNP complex represents an intermediate particle in U5 snRNP biogenesis (Gottschalk et al. 2001). Also, Spp382p was associated with Prp8T/770p but was not detectable with full-length Prp8p or Prp8T/2173p. Spp382p has been affinity-selected previously with Prp8p and other splicing factors (Gavin et al. 2002) but is not known to be a stable component of U5 snRNPs. Thus, Prp8T/770p produced a different profile of coselected proteins compared to full-length Prp8p and this likely includes an Aar2p-U5 snRNP precursor complex related to that described by Gottschalk et al. (2001).

As Western blotting indicated that the N-terminal fragments (1–770 and 1–2173) were detectable in the TAP-purified fraction, albeit at a low level (data not shown), the effects of different salt concentrations were tested in order to establish conditions that gave complete separation of the N- and the C-terminal Prp8p fragments. The Prp8p C-terminal TAP-tagged fragment was affinity-selected using IgG agarose beads, and the presence of the Prp8p N-terminal fragment in coprecipitates was investigated by Western blotting with anti-8.6 antibodies (which recognize the N-terminal Prp8 polypeptides (Fig. 3A), and these immune precipitates were analyzed by Western blotting with anti-8.6 antibodies to pull down the N-terminal Prp8 polypeptides. This single-step affinity selection in the presence of 150 mM or 300 mM salt did not disrupt the association of the N- and the C-terminal Prp8p fragments; however, with 500 mM salt, only C-terminal TAP-tagged Prp8p polypeptides were selected (Fig. 2B). It should be noted that 500 mM salt also disrupted the association of the U5 snRNA with Prp8p (data not shown).

**Association of Snu114p and Aar2p with Prp8p**

The association of N- and C-terminal fragments of Prp8p with other splicing factors was investigated in the presence of 500 mM NaCl. TAP-tagged C-terminal Prp8p fragments were pulled down with IgG-agarose, and the supernatant were then incubated with anti-8.6 antibodies to pull down the N-terminal Prp8 polypeptides (Fig. 3A), and these immune precipitates were analyzed by Western blotting for the presence of Prp8p N-terminal fragments, Snu114p, or Aar2p (Fig. 3).

Snu114p was detected associated with TAP-tagged full-length Prp8p (Fig. 3B, lane 1) and with C-terminal Prp8p 437–2413 (Fig. 3B, lane 3) but not with Prp8p 771–2413 or 2174–2413 (Fig. 3B, lanes 4, 5). Following the precipitation with anti-8.6 antibodies, Snu114p was found associated with non-TAP-tagged Prp8p (Fig. 3B, lane 7) and with N-terminal polypeptides 1–770 and 1–2173 of Prp8p (Fig. 3B, lanes 9, 10) but not with full-length TAP-tagged Prp8p or with amino acids 1–439 (Fig. 3B, lanes 6, 8), as most Snu114p had already been pulled out of these samples during the first-round precipitation of the TAP-tagged
C-terminal polypeptides (Fig. 3B, lanes 1,3). Thus, Snu114p associates with the region 437–770 of Prp8p.

Aar2p was present in precipitates with TAP-tagged full-length Prp8p and with Prp8p residues 437–2413 and 771–2413 (Fig. 3C, lanes 1,3,4) but not with untagged Prp8p, or with Prp8p amino acids 2174–2413 (Fig. 3C, lanes 2,5). Even when the Prp8T/2173p immunoprecipitation was performed under conditions (150 mM NaCl) that did not disrupt the association of the N- and the C-terminal Prp8p fragments (or of the U5 snRNA), Aar2p was not coprecipitated (Fig. 3D, lane 7). In addition, centrifugation through a 10%–30% glycerol gradient (100 mM KCl) showed that Aar2p mostly cofractionates with full-length Prp8p and Snu114p (Fig. 4A, lanes 5–11), although some Aar2p was also detected near the top of the gradient (Fig. 4A, lane 25) as was observed previously (Gottschalk et al. 2001). In contrast, the Aar2p in Prp8T/2173 extract was mostly found in the lower density fraction (Fig. 4B, lane 25). Thus, the interaction of Aar2p with Prp8p was disrupted by dissecting Prp8p at position 2173.

**DISCUSSION**

Yeast Prp8p is highly conserved in both its sequence (61% identical to human) and its large size. Bioinformatic approaches predicted an MPN (Mpr-1, Pad-1, N-terminal) domain, defined by amino acids 2173–2310, an RRM (RNA Recognition Motif) in the region 1059–1151, and potential NLS(nuclear localization signals, residues 95–156 (Grainger and Beggs 2005), but there is no other obvious structural domain within this huge protein. By functional, genetic classification, Kuhn et al. (Kuhn and Brow 2000; Kuhn et al. 2002) designated five regions: a (residues 233–365), b (residues 608–687), c (residues 785–864), d (residues 1091–1197), and e (residues 1621–1878).

Figure 5 illustrates the sites of the viable disruptions that we identified relative to these other known features of Prp8p. Considering (1) that a very large number of disrupted prp8 alleles were screened for viability (in excess of 40,000 within the PRP8 open reading frame), (2) the broad spread of insertions observed by DNA digest analysis, and (3) the evidence that Tn5 insertions occur essentially randomly (Shevchenko et al. 2002), it seems unlikely that the three regions of PRP8 in which viable insertions were isolated might represent hotspots for Tn5 insertion. This is further supported by the fact that many viable insertions were found in the nonessential N-terminal proline-rich region of Prp8p. Therefore, the failure to obtain viable constructs with transposon insertions in the 771–2173 region of Prp8p suggests that the genetically defined regions c, d, and e and the RRM function as one indivisible unit. Thus the results presented here show that Prp8p can be dissected into four functional domains: domain I, residues 1–394; domain II, 443–770; domain III, 771–2170; and
domain IV, 2179–2413. The dissection points near the C terminus of Prp8p (2170–2179) lie precisely at one boundary of the predicted MPN/JAB domain (2173–2310). This is consistent with the notion that these transposon insertions may be tolerated because they have targeted a “linker” sequence upstream of this functional module. Furthermore, the predicted NLS and the genetically defined region a lie within domain I, region b is within domain II, and the RRM and regions c, d, and e are within domain III.

To investigate which other splicing factors interact with the different domains of Prp8p, 500 mM salt was used to disrupt the associations of the N- and the C-terminal Prp8 polypeptides. The need for such stringent conditions to separate the pairs of dissected Prp8 polypeptides shows that strong intra- and/or intermolecular interactions hold the two ends of the protein together. Kuhn and Brow (2000) proposed the existence of a coiled-coil interaction between residues 640–670 and 1625–1650 and, if this physical interaction exists, it might be responsible for reconstituting the Prp8p molecule that is split at 770/771.

The strong association of the pairs of Prp8 polypeptides limited their utility as a tool to investigate which domains of Prp8p interact with other splicing factors, as 500 mM salt also disrupted the association of Prp8p with US snRNA. Nevertheless, using a two-step immunoprecipitation strategy (Fig. 3), the association of Snu114p with Prp8p was found to require Prp8p residues 437–770, which is in close agreement with two-hybrid interaction data in which Snu114p as bait selected multiple Prp8p fragments that had amino acids 420–464 in common (I. Dix and J.D. Beggs, unpubl.). The strong association between Prp8p and Snu114p is not surprising, as the human homologs (h220K and h16K) have been identified in an RNA-free complex and remain associated in 0.4 M of the chaotropic salt sodium thiocyanate (Achsel et al. 1998).

Interestingly, even at moderate salt concentrations (150 mM NaCl or 100 mM KCl), the association of Aar2p was disrupted by dissecting Prp8p at residue 2173. Previously, Gottschalk et al. (2001) purified yeast Aar2p associated with Prp8p and Snu114p. Here we show that Prp8p/771–2413p could coselect Aar2p in the absence of Snu114p (Fig. 3C, lane 4), and Prp8p/2173p could coselect Snu114p in the absence of Aar2p (Figs. 3B [lane 10], 4B). This suggests that the associations of Snu114p and Aar2p with Prp8p are not interdependent.

In summary, we describe here a novel approach to characterize the functional domains of a protein in vivo and demonstrate its utility by identifying four functional domains within the very large yeast Prp8p splicing factor, one of which—domain III—appears to require a continuous stretch of 1400 amino acids. Although Prp8p has been demonstrated to function when expressed as pairs of trans-acting polypeptides, each of these disruptions was detrimental to growth; the 2173 dissection caused slower growth and the 436 and 770 dissections resulted in heat sensitivity. For the 2173 dissection at least, the integrity of Prp8p’s interaction with another protein was destabilized. Therefore, this approach has also confirmed the functional significance of the very large size of this highly conserved protein that has been proposed to function as a scaffold for molecular interactions in the spliceosome.

**MATERIALS AND METHODS**

Plasmids and yeast strains used in this work are listed in Table 1 and oligonucleotides are listed in Table 2.

**Transposition methods**

The gene dissection transposon was assembled in the vector pMOD(MCS) (Epigen) by inserting three PCR products as follows: (1) PRP8 gene downstream flanking region (EcoRI site), (2) Kanamycin-resistance cassette (BamHI site), and (3) PRP8 gene upstream flanking region (HindIII site). An SV40 NLS tag was then inserted into the unique AflII site near the downstream end of the transposon. The structure of the transposon in the resulting pMOD-Tn plasmid is therefore as follows: ME PRP8 gene downstream flanking region Kan PRP8 gene upstream flanking region ATG-NLS ME where “ME” denotes the Mosaic End sequences recognized by Tn5 Transposase (Fig. 1). Transposon DNA was released from the pMOD-Tn plasmid for in vitro transposition reactions by cleavage with PvuII and gel purification. For random insertion of the transposon into the entire PRP8 ORF the target DNA was pRS314 carrying a 9.6-kb fragment of yeast genomic DNA including the PRP8 open reading frame and flanking sequences with unique Nhel and MluI sites introduced at the N and the C termini of the coding sequence, respectively. An SV40 NLS tag was also introduced at the Nhel site (oligonucleotides AW40 and AW41) prior to the transposition step so that both Prp8p fragments can be localized to the nucleus after gene dissection. Target DNA (1 μg) was incubated in a 10-μL reaction with a molar equivalent of the transposon in the presence of 0.1 μL EZ::TN transposase (Epigen) at 37°C for 2 h, according to the manufacturer’s instructions. The reaction was stopped by addition of 1 μL 1% SDS followed by

---

**TABLE 1. Plasmids and yeast strains**

<table>
<thead>
<tr>
<th>Name</th>
<th>Details of plasmid/genotypes of yeast strains</th>
</tr>
</thead>
<tbody>
<tr>
<td>pMod-Tn</td>
<td>PRP8 transposon plasmid. Contains PRP8 downstream flanking region, KanMX4 cassette, PRP8 upstream flanking region, SV40 NLS.</td>
</tr>
<tr>
<td>pBS1479</td>
<td>PCR template for TAP-tagging (Rigaut et al. 1999), SV40 NLS.</td>
</tr>
<tr>
<td>SC261ΔB1</td>
<td>MATa, ura3-52, leu2, trp1, pep4-3, prb1-1132, prc1-407, prp8::Blasticidin [pRS316:PRP8]</td>
</tr>
<tr>
<td>BMA38a</td>
<td>MATa, his3A200, leu2-3,112, ura3-1, trp1Δ, ade2-1, can1-100</td>
</tr>
<tr>
<td>RG8T</td>
<td>PRP8::TAP-TRP1, otherwise as BMA38a</td>
</tr>
<tr>
<td>RB8T/2173</td>
<td>PRP8(1-2173);Tn:PRP8(2174-2413);TAP-TRP1, otherwise as BMA38a</td>
</tr>
<tr>
<td>KL8T/770</td>
<td>PRP8(1-770);Tn:PRP8(771-2413);TAP-TRP1, otherwise as BMA38a</td>
</tr>
<tr>
<td>KL8TG/436</td>
<td>His3::P_GAL1-HA3-P1:PRP8(1-439);Tn:PRP8(437-2413);TAP-TRP1, otherwise as BMA38a</td>
</tr>
</tbody>
</table>
TABLE 2. Oligonucleotides

<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AW40 and AW41 introduce an SV40 NLS at Nhel site (N terminus of Prp8p)</td>
<td></td>
</tr>
<tr>
<td>AW40 CTAGCGACCTCCAAAAAGAAAGAGAAATGACTGCTGATAGGCTGATAGGTGTCAGGTGGAGTTG</td>
<td></td>
</tr>
<tr>
<td>AW41 CTAGCGACCTCCAAAAAGAAAGAGAAATGACTGCTGATAGGCTGATAGGTGTCAGGTGGAGTTG</td>
<td></td>
</tr>
<tr>
<td>AW22 TGCTGGATCGATGGATGAGAGATTCTGACTGACTGCTGATAAGGATGCAAGGATGCAAGGATGCAAGG</td>
<td></td>
</tr>
<tr>
<td>AW25 and AW26 for Kan gene PCR (BamHI)</td>
<td></td>
</tr>
</tbody>
</table>
| AW25 GCCAGCAGATCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA

heating to 70°C for 10 min. Aliquots (1 µL) of the transposition reaction were transformed by electroporation into XL10 Gold cells (Stratagene), and plasmids that had been targeted by the transposon were recovered by selection for ampicillin and kanamycin resistance. Plasmid DNA was prepared from the pooled transformants, and suitable fragments of the prp8 ORF containing transposons were excised, gel purified, and recloned into "gapped" pRS314-PRP8 vector, again selecting for ampicillin and kanamycin resistance. Plasmid DNA was prepared from the pooled transformants to produce a library in which each individual carries a transposon inserted in a region of the prp8 ORF. Functional copies of PRP8 were identified by plasmid shuffle after transformation into haploid yeast strain SC261AB1 in which the genomic copy of PRP8 is deleted and survival is dependent on a copy of PRP8 on pRS316 (marked with URA3). Cells harboring a functional transposon-tagged copy of PRP8 were isolated by growth on plates containing 5-fluoroorotic acid (5-FOA) to select for loss of the URA3-marked plasmid copy of PRP8. Plasmid DNA was isolated from these cells and retested by plasmid shuffle. The location of the transposon was determined approximately by restriction mapping and confirmed by DNA sequencing. Note that the transposition event results in the duplication of three amino acids at the site of insertion; e.g., insertion after codon 2170 results in the N-terminal polypeptide terminating with amino acid 2173 and the other starting with amino acid 2171.

Construction of yeast strains with dissected chromosomal prp8 alleles

Strain RG8T was made by integration of a linear DNA cassette, generated by PCR, that encodes the TAP tag sequence (from plasmid bRS1479) (Rigaut et al. 1999), fused in-frame with the 3' end of the PRP8 open reading frame, plus the TRP1 gene as selectable marker. For insertion of pseudotransposons (mosaic regions omitted) into this chromosomal TAP-tagged copy of PRP8 the Kanamycin-resistance marker was replaced with a KanMX4 cassette (used for dissection after codon 436) (Wach et al. 1994), or, alternatively, a copy of the URA3 gene from Kluveromyces lactis was inserted between the unique XbaI and PstI sites downstream of the Kanamycin-resistance marker (used for dissection after codons 770 and 2713) and PCR-generated DNA was used for yeast transformation, inserting at specific positions in the TAP-tagged chromosomal copy of PRP8 by homologous recombination. For dissection at position 436, the PCR product was generated with primers in the flanking PRP8 sequences. For dissection at positions 770 and 2173, flanking PRP8 sequences were included in the PCR primers and added directly to the transposon downstream and upstream flanking elements (Table 2). G418 (geneticin) resistant or Ura⁺ transformants were selected, correct integration of the cassettes was checked by PCR, and production of the TAP-tagged Prp8p was monitored by Western blotting.

Purification of Prp8p complexes by TAP

Yeast whole-cell extracts were produced (Umen and Guthrie 1995) from 30 L of log phase cells, and tagged proteins were isolated (Rigaut et al. 1999). The purified proteins were resolved on a Novex NuPAGE 4%–12% bis-Tris gel (Invitrogen), using morpholinepropanesulfonic acid (MOPS) running buffer and stained with Sypro ruby (Biorad), and the stained bands were subjected to mass spectrometric analysis by the “FingerPrints” Proteomics Facility (Wellcome Trust Biocentre, University of Dundee) and EPIC Proteomics service (Wellcome Trust Centre for Cell Biology, University of Edinburgh).

Immunoprecipitation and immunoblotting

Immunoprecipitations were performed (Teigelkamp et al. 1995) using IgG-agarose (Sigma) or anti-8.6 Prp8 polyclonal antibodies (raised against an N-terminal peptide) in 500 mM NaCl unless stated otherwise. Proteins were resolved in Novex NuPAGE 4%–12% bis-Tris gels (Invitrogen), electroblotted to a polyvinylidene difluoride membrane (Hybond-P, Amershams Bioscience), probed with anti-Snu II 4p antibodies (Bartels et al. 2003) or anti-Aar2p antibodies (Gottschalk et al. 2001), and detected using horseradish peroxidase-conjugated goat anti-rabbit antibodies and ECL reagents (Amershams Bioscience).
Glycerol gradient analysis

Glycerol gradient analysis was performed (10%–30% glycerol gradient containing GG buffer; Bartels et al. 2003) using 80 µL splicing extract diluted with 120 µL GG buffer (20 mM HEPES at pH 7.0, 100 mM KCl, 0.2 mM EDTA; final glycerol concentration 8%). After sedimentation at 37,000 rpm for 17 h in a SW40 Ti rotor (Beckman) at 4°C, 400-µL fractions were collected and stored at −70°C. Alternate fractions were phenol/chloroform extracted, and proteins were acetone precipitated and assayed by Western blotting.

ACKNOWLEDGMENTS

We thank Dr. Patrizia Fabrizio for the generous gift of anti-Aar2p and anti-Snu114p antibodies and Dr. Duncan Short for mass spectrometric analysis of the Aar2 and Lsm2 proteins. This work was funded by a studentship from The Darwin Trust of Edinburgh to K.-L.B., by Wellcome Trust Grant 067311 to J.D.B., and by Medical Research Council funding to A.J.N., and benefited from the facilities of the Edinburgh Protein Interaction Centre. J.D.B. is Royal Society Darwin Trust Research Professor.

Received October 27, 2005; accepted November 10, 2005.

REFERENCES


Umen, J.G. and Guthrie, C. 1995. Ppr16p, Slu7p, and Ppr8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. RNA 1: 584–597.
