# Detection and Analyses of Proteins Interacting with the U6 snRNA in SnRNPs and in Spliceosomes in the Yeast *Saccharomyces cerevisiae.*

**Valérie Paulette Incarnacion VIDAL**

**A Thesis Presented for the Degree of Doctor of Philosophy**

**University of Edinburgh**

**July 1998**

## Declaration

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

## Acknowledgements

Special thanks to Jean for giving me the opportunity to work in her lab and supporting me during this research project. Also many thanks for all the time she spent correcting my thesis and teaching me how to write in English. I would also like to thank every member of the Beggs group, past and present, for providing such a good working environment and for being patient and trying to understand my Johnny Foreigner language. Many thanks for Margaret, Alan, Paul and my flatmates for reading through my horrible chapters and correcting my English over and over again. Hey folks... that's the end of your torture!!

Also thanks to my "Banana brother" who kept producing every week a new Sm-like fusion protein.

Lastly, I am indebted to my parents for supporting me over the years to enable me to get this far. Also thanks for my friends who helped me a lot during this hard time.

My last wishes would be that a nice soul from the lab (could be more than one actually) would look after the fish, feed them and clean the fish tank when needed. Also could somebody nag after the boys to tidy up their mess.....

# ABSTRACT

The yeast U6 snRNA is the most highly conserved of the snRNAs involved in nuclear pre-mRNA splicing. U6 snRNA is a very important and dynamic molecule, undergoing several conformational changes during spliceosome assembly. In snRNP particles and in the spliceosome, U6 snRNA is involved in numerous well characterised RNA-RNA interactions. In contrast, there have been few studies of the U6 snRNA-protein interactions. With the exception of the Prp24p splicing factor, no direct U6 snRNA-protein binding has been reported to date.

An experimental procedure was developed to detect proteins associated with the U6 snRNA: In a yeast nuclear extract, functional U6 snRNP particles were reconstituted with a radiolabelled U6 snRNA produced by transcription *in vitro*. The extract was UV-irradiated to covalently crosslink proteins that are in close proximity to the U6 snRNA. Proteins were subsequently immunoprecipitated, either directly or via snRNP particles. After separation on SDS-PAGE, proteins UV-crosslinked to radiolabelled U6 snRNA were detected by autoradiography. Using this procedure, the splicing factor Prp24p, and the U6 snRNP-associated proteins Uss1, Uss4 and Uss5 were found to be in contact with the U6 snRNA in the U6 snRNP, the U4/U6.U5 tri-snRNP and probably in the U4/U6 snRNP. Other (uncharacterized) proteins were also UV-crosslinked to the U6 snRNA in these particles. The U5 snRNP specific Prp8 protein interacts with the U6 snRNA in the tri-snRNP. The evidence that Prp8p contacts the U6 snRNA in this particle is in keeping with the proposed role for Prp8p stabilising the U4/U6 snRNA interaction in the tri-snRNP. Further investigations showed that the U6 snRNP-specific protein Spb8 binds the U6 snRNA in an unspecific manner while Uss6p and Uss3p interact weakly with this RNA molecule. No UV-crosslinking was detected between Uss2p or Uss7p and the U6 snRNA.

Analysis of protein-U6 snRNA contacts occuring in the spliceosome indicates that two unidentified polypeptides of 30 kD and 40 kD are in close proximity to the U6 snRNA,

at least before the first step of the splicing reaction. No contact between $Prp2^{LAT}$ mutant protein and U6 snRNA could be detected in spliceosomes.

The effect of U6 snRNA mutations on these RNA-protein interactions was also investigated. The most significant effects were obtained for the C48A and G60C mutations and for a *Fok*I truncation (lacking the last 18 nucleotides of U6 snRNA) which have differential effects on the crosslinking of Prp24p, Prp8p and Uss1p to U6 snRNA. The implication of the *Fok*I truncation on protein interactions (in particular Prp24p and Uss1p contacts) is discussed. The UV-crosslinking data suggest that the last 18 nucleotides of the U6 snRNA may be a binding site for Prp24p and Uss1p and/or it may promote the binding of these proteins elsewhere in this snRNA.

U6 snRNP particles were also reconstituted with U6 snRNA uniquely radiolabelled at position $U_{54}$. UV-crosslinking experiments demonstrated that Prp8p binds strongly to U6 snRNA at this important position. In comparison, Uss1p interacts much less strongly, suggesting that $U_{54}$ is not the main contact point of this protein.

The results presented here form an initial characterisation of the protein contacts made with the U6 snRNA in snRNP complexes. Interestingly, the binding of the U6 snRNP-specific proteins Uss1, Uss4 and Uss5 to the U6 snRNA in the U6, U4/U6 snRNPs and in the tri-snRNP suggests that these polypeptides may have an important role to play. These polypeptides share significant homology with the Sm proteins and, by analogy, it can be speculated that they may be involved in U6 snRNP biogenesis.

# COMMON ABBREVIATIONS

| | |
|---|---|
| **%** | Percentage |
| **A** | Ampere |
| **Amp** | Ampicillin |
| **ARS** | Autonomous Replication Sequence |
| **ATP** | Adenosine 5'-triphosphate |
| **BCIP** | 5-Bromo-4-Chloro-3-Indolyl Phosphate |
| **bisacrylamide** | *N,N'*-methylene-bisacrylamide |
| **bp** | Base Pair |
| **BSA** | Bovine Serum Albumin |
| **°C** | Degree Celsius |
| **cDNA** | Complementary Deoxyribonucleic Acid |
| **CEN** | Centromere sequence |
| **cfu** | Colony forming unit |
| **Ci** | Curie |
| **cpm** | Count per minute |
| **CTP** | Cytidine 5'-triphosphate |
| **D** | Dalton |
| **dATP** | Deoxyadenosine triphosphate |
| **dCTP** | Deoxycytidine triphosphate |
| **ddATP** | Dideoxyadenosine triphosphate |
| **ddCTP** | Dideoxycytidine triphosphate |
| **ddGTP** | Dideoxyguanosine triphosphate |
| **ddTTP** | Dideoxythymidine triphosphate |
| ***DNA*** | Deoxyribonucleic acid |
| **DNAse** | Deoxyribonuclease |
| **dNTP** | Deoxynucleoside triphosphate |
| **ds** | Double-stranded |
| **DTT** | Dithiothreitol |
| **dTTP** | Deoxythymidine triphosphate |
| **dUTP** | Deoxyuridine triphosphate |
| **EDTA** | Ethylenediaminetetraacetic acid |
| **Fig.** | Figure |
| **fmole** | Femtomole |
| **g** | Gram(s) |
| **Gal** | Galactose |
| **Gal4** | Gal4 DNA binding domain |
| **Glu** | Glucose |
| **GST** | Glutathione S-Transferase |
| **h(s)** | Hour(s) |
| **HA** | Hemagglutinin |
| **HEPES** | *N*-2-hydroxyethylpiperazine-*N'*2-ethanesulfonic acid |
| **IAA** | 3-β indoleacrylic acid |
| **IP** | Immunoprecipitation |

| | |
|---|---|
| **IPTG** | Isopropyl-β-*D*-thiogalactopyranoside |
| **Kb** | Kilobase |
| **kD** | KiloDalton |
| **L** | Litre(s) |
| **Lab** | Laboratory |
| **LB** | Luria Broth |
| **LexA** | LexA DNA binding domain |
| **M** | Molar |
| **m3G** | 2, 2, 7-trimethyl guanosine |
| **mA** | Milliampere |
| **mCi** | Millicurie |
| **mg** | Milligram(s) |
| **min** | Minute(s) |
| **ml** | Millilitre(s) |
| **mm** | Millimetre(s) |
| **mM** | Millimolar |
| **MOPS** | 3-(*N*-morpholino)propane-sulfonic acid |
| **NBT** | Nitroblue tetrazolium |
| **nM** | Nanomolar |
| **nmole** | Nanomole |
| **NP40** | Nonidet P-40 detergent |
| **nt** | Nucleotide |
| **NTP** | Nucleoside triphosphate |
| **OD** | Optical Density |
| **OLB** | Oligonucleotide Labelling Buffer |
| **ONPG** | *O*-nitrophenyl-β-galactosidase |
| **ORF** | Open Reading Frame |
| **P** | Plasmid designation |
| **PAB** | Poly(A) binding |
| **PAGE** | Polyacrylamide Gel Electrophoresis |
| **PBS** | Phosphate Buffered Saline |
| **PCR** | Polymerase Chain Reaction |
| **PEG** | Polyethylene Glycol |
| **pfu** | Plaque forming unit |
| **pmole** | Picomole |
| **PNK** | Polynucleotide Kinase |
| **R** | Resistant |
| **RNA** | Ribonucleic Acid |
| **RNase** | Ribonuclease |
| **RNasin** | Ribonuclease Inhibitor |
| **rpm** | Rotation per minute |
| **rRNA** | Ribosomal Ribonucleic Acid |
| **RT** | Room Temperature |
| **sec** | Second(s) |
| **SDS** | Sodium Dodecyl Sulfate |
| **snRNA** | Small Nuclear Ribonucleic Acid |

| | |
|---|---|
| **snRNP** | Small Nuclear Ribonucleoprotein Particle |
| **TEMED** | *N,N,N',N'*-tetramethyl-ethylenediamine |
| **TMG** | Trimethylguanosine |
| **tRNA** | Transfer Ribonucleic Acid |
| **trp** | Tryptophan |
| **U** | Unit(s) |
| **UpG** | Uridylyl(3'-5')Guanosine |
| **UV** | Ultra Violet |
| **μCi** | Microcurie |
| **μg** | Microgram(s) |
| **μl** | Microlitre(s) |
| **μM** | Micromolar |
| **v** | Volume |
| **v/v** | Volume per unit volume |
| **V** | Volts |
| **w/v** | Weight per unit volume |
| **w** | Weight |
| **w/o** | Without |
| **X** | Fold |
| **Xgal** | 5-Bromo-4-Chloro-3-Indolyl-β-D-Galactoside |

**Oligo knockout:** Oligodeoxynucleotide complementary to a particular snRNA sequence (Table 2.5). It was added in a yeast cellular extract in order to target the degradation of a particular snRNA via endogenous RNase H activity (sections 2.6.2.3-4).

**U6 RNA:** U6 snRNA produced by transcription *in vitro* with the T7 RNA polymerase (section 2.4.4.1).

**U6 RNA ($^{32}$P-αATP):** U6 snRNA radiolabelled with $^{32}$P-αATP. This RNA was produced by transcription *in vitro* using the linearised pUC12U6 plasmid DNA (Table 9) as template.

**U6 RNA ($^{32}$P-αUTP):** U6 snRNA radiolabelled with $^{32}$P-αUTP. This RNA was produced by transcription *in vitro* using the PUC6 DNA template amplified by PCR reaction with oligodeoxyribonucleotides Lory1 and S5761 (Table 8) as primers.

## Amino Acids (single letter code)

| | | | |
|---|---|---|---|
| A | Alanine | M | Methionine |
| C | Cysteine | N | Asparagine |
| D | Aspartate | P | Proline |
| E | Glutamate | Q | Glutamine |
| F | Phenylalanine | R | Arginine |
| G | Glycine | S | Serine |
| H | Histidine | T | Threonine |
| I | Isoleucine | V | Valine |
| K | Lysine | W | Tryptophan |
| L | Leucine | Y | Tyrosine |

# TABLE OF CONTENTS

## *Chapter One - Introduction*

## *Chapter Two - Materials and Methods*

## *Chapter Four - The U6 SnRNP Reconstitution*



## *Chapter Five - Detection and Characterisation of proteins Interacting with the U6 SnRNA*



## *Chapter Six - Investigation of the Direct Binding of the Sm-like Proteins to the U6 SnRNA*

# CHAPTER ONE

## Introduction

# INTRODUCTION

By the late 1970s, the physical structure of genes was firmly established from work in bacteria; the sequences of the gene, the RNA and the protein were colinearly organised and expressed. Since the science of genetics suggested that genes in eukaryotic organisms behaved similarly to these of prokaryotic organisms, it was then assumed that the bacterial gene structure was universal. However, further analyses of the genetic material in eukaryotic cells, indicated that the genes were physically separated by the nucleus' membrane from the messenger RNAs translated in the cytoplasm. This observation suggested that the molecular biology of gene expression in eukaryotes was not as straightforward as that in bacteria. In 1977, the study of the adenovirus expression led to the discovery of split genes and precursor-mRNA splicing (Berget *et al.*, 1977; Chow *et al.*, 1977). Shortly afterwards a number of cellular genes were shown to be interrupted by intervening sequences or introns. Amongst the first intron-containing genes discovered, the globin gene (Jeffreys and Flavell, 1977; Tilghman *et al.*, 1978) and the ovalbumin gene (Breathnach *et al.*, 1977) were shown to contain respectively two and eight introns. Further studies indicated that most eukaryotic genes harboured introns. In 1981, Breathnach and Chambon demonstrated that these introns possessed limited conserved sequences at their boundaries. Interestingly, these conserved sequences were common to plant, yeast and vertebrate cells (Padgett *et al.*, 1986), suggesting that the splicing process was evolutionarily conserved. Since then, the splicing mechanism has been widely studied in various organisms highlighting the complexity of this process.

Introns have been classified according to their conserved sequences, structures and mechanism of splicing. Six major groups of introns were distinguished: group I; group II; group III; tRNA introns; nuclear introns removed by the U2-type spliceosome and nuclear introns excised by the U12-type spliceosome. The removal of these different introns involves different mechanisms which share several similarities, supporting the idea of evolutionary conservation (for review see Lambowitz and Belfort, 1993).

A brief presentation of the group I, II and III introns and their splicing mechanisms will be related in this chapter, highlighting the similarities of their structures and splicing processes. A more detailed description of the nuclear introns removed by the U2-type spliceosome and their splicing mechanism will be given. This will be followed by a short description of the U12-type splicing machinery.

As nuclear introns are distributed in a wide range of organisms, their characterisation will be limited to that of introns located in the nuclear pre-mRNAs in mammals and in the yeast *Saccharomyces cerevisiae*.

## 1.1 Group I Introns

Group I introns possess the widest phylogenetic distribution and are found in both eubacteria and eukaryotes (although not in vertebrates). They interrupt genes coding for proteins, rRNAs and tRNAs located in mitochondria, chloroplast and nuclear genomes. Group I introns are sporadically distributed, and can be present in one strain or species but absent in another of the same genus. Their discovery in bacteriophage T4 marked the first time that RNA splicing was found in prokaryotes (Chu *et al.*, 1986; Gott *et al.*, 1986).

***Common features:*** different group I introns have limited sequence homology but share conserved secondary and tertiary structures. The secondary structure of group I introns is shown in Figure 1A, and consists of a combination of short- and long- range pairings denoted P1 through P10, with elements P1-P8 forming the catalytic core. Most group I introns include an "internal guide sequence" (IGS) which is shown in Figure 1B. This structure holds the exons together and ensures that exon ligation occurs. The presence of a divalent metal ion is critical for group I intron structure and function. $Mg^{2+}$ or $Mn^{2+}$ are required for intron' activity while $Ca^{2+}$, $Sr^{2+}$ and $Ba^{2+}$ can mediate the correct overall folding of the RNA but cannot confer catalytic activity. This suggests that two classes of metal ion-binding sites (structural and catalytic) must be occupied (Cech,

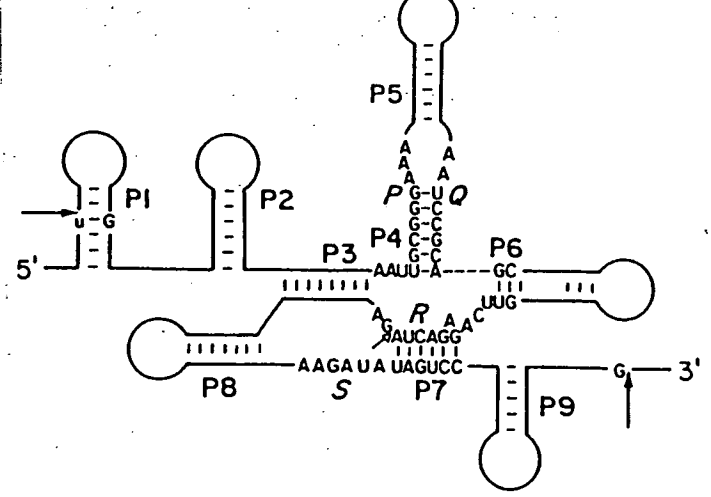
**Figure 1. Group I introns**

A: Secondary structure of the group I intron of Tetrahymena precursor rRNA (Michel *et al.*, 1990). This diagram shows the conserved nucleotides and the secondary structure of the group I intron. The paired structures are shown (P1-P9). Two arrows indicate the conserved residues located at the 3' end of the exon and the 3' end of the intron. An opened arrow indicates a G residue within an AGA motif of the intron (residue $G_{264}$ of the P7 domain of the Tetrahymena intron) (Michel *et al.*, 1989). This guanosine ($G_{264}$) binds to a free guanosine in the first step of splicing (see Figure 1C), and also serves to recognize the 3' terminal guanosine of the intron during the second step of splicing (Been *et al.*, 1991). The most conserved nucleotides contained within four short sequences are annoted as P, Q, R and S. The distance between these sequence elements can vary from ~20 nt to many hundreds.

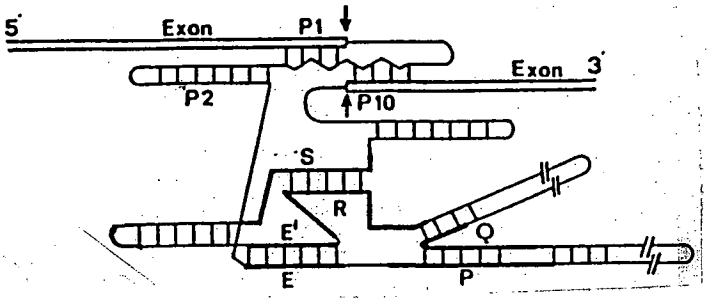
**B:** Diagrammatic representation of the secondary structure of a group I intron. This scheme was adapted from Davies *et al.* (1982). The paired structures and the conserved sequences (P, Q, R, S) are shown. On this particular diagram, the "internal guide sequence" is indicated by a wavy lane. This sequence pairs at the end of the upstream exon (P1) and at the beginning of the downstream exon (P10) acting as an adaptor molecule that brings the two splice sites tightly together.

**C:** Schematic representation of the group I intron splicing reaction. Almost invariantly the conserved residues at the 3' end of the exon and at the 3' end of the intron are respectively U and G. The first step of splicing requires the specific binding of a free guanosine. The 3' OH group of this free guanosine attacks (nucleophilic attack) the 3' phosphodiester bond of the 5' exon-intron junction. Thus after cleavage of the 5' exon-intron junction, the free G becomes covalently attached to the 5' end of the intron. In the second step of the reaction, the terminal G residue of the intron is attacked by the terminal 3' OH of the 5' exon leading to exon ligation. With some group I introns, cyclisation occurs, coupled to cleavage and release of a short 5' oligonucleotide harboring the G residue at the 5'-terminal.

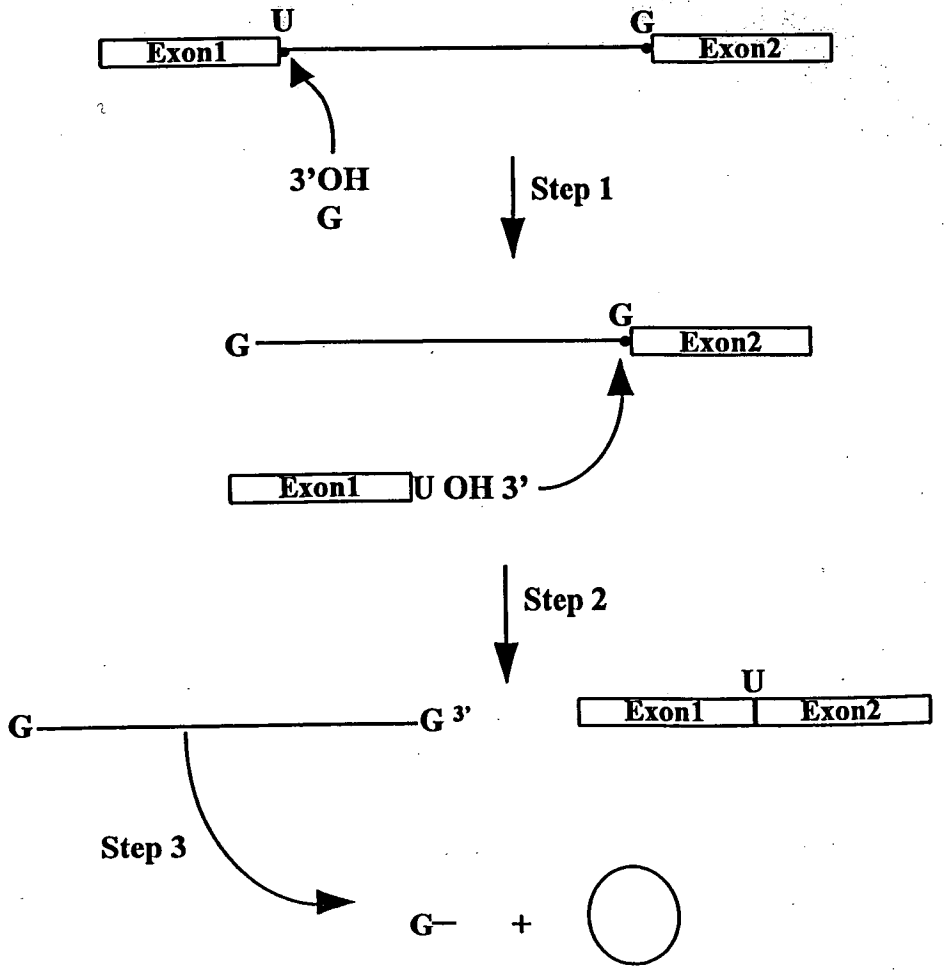
A


P5
P1
P2
P3
P4
P6
P8
S
P7
P9
5'
3'

B


5'
Exon
P1
Exon 3'
P2
P10
S
R
E'
Q
E
P

C


U
G
Exon1
Exon2
3'OH
G
Step 1
G
G
Exon2
Exon1
U OH 3'
Step 2
U
G
G 3'
Exon1
Exon2
Step 3
G—
+

1993). Group I introns have only a single catalytic centre and undergo splicing by a two-step trans-esterification mechanism which is shown in Figure 1C. Although some group I introns self-splice *in vitro*, many require proteins for efficient splicing *in vivo* (Lazowska *et al.*, 1989). Group I introns vary in size from 200 nt to 3000 nt, depending on the length of the peripheral sequences and whether or not they contain an open reading frame (ORF). Some of these ORFs encode essential cellular proteins (e.g. mitochondrial ribosomal protein S-5, which is encoded within the *Neurospora* mitochondrial large rRNA intron; Burke and RajBhandary, 1982), the majority of them function as maturases associated with a larger family of site-specific endonucleases. These endonucleases target intronless alleles of the intron-containing genes, mediating intron mobility.

***Similarities with other groups of introns:*** in Tetrahymena, mutations in the invariant $AG_{264}A$ motif of group I introns have similar effects on the second step of splicing, to comparable mutations in the invariant AGA motif of the U2 snRNA in *S. cerevisiae* (McPheeters and Abelson, 1992). Moreover, these two motifs are located in similar structural environments. In group I introns, the AGA sequence lies within the P7 helix of the catalytic core structure. In the U2 snRNA this motif lies within helix I, constituted by U2 snRNA and U6 snRNA base-pairing and located in the catalytic centre of the spliceosome.

## 1.2 Group II Introns

Group II introns exist in relatively small number in fungal mitochondrial genomes. They are predominant in plant mitochondria and chloroplasts, are present in algae and overwhelmingly frequent in *Euglena gracilis* chloroplasts.

***Common features:*** group II introns are multidomain RNAs endowed with catalytic (ribozyme) activity by which they direct and catalyse the splicing of neighbouring exon sequences. Group II introns that can direct a complete self-splicing reaction in the

absence of protein do so only under non-physiological conditions, and depend on proteins for *in vivo* activity. Only a minority of group II introns are proven ribozymes. A number of group II introns encode proteins of the reverse transcriptase family (maturases), which enable them to change genomic location. Such introns are bifunctional and considered as active participants of the splicing reaction, as well as mobile genetic elements. The secondary structure of group II introns is shown in Figure 2A. It comprises a central wheel from which radiate six spokes that define the six major double-helical domains (I to VI). Two major subdivisions (IIA and IIB) of group II introns are distinguished, based on subtle structural and sequence differences (Michel *et al.*, 1989b). Splicing of group II introns proceeds via a two-step trans-esterification mechanism which is shown in Figure 2B. This mechanism is essentially identical to that involved in nuclear pre-mRNA splicing (see section  for description) (Lambowitz and Belfort, 1993). The splicing of group II introns involves a single active site which rearranges its components to allow the second step to occur (Chanfreau *et al.*, 1996). In particular, structural elements located in peripheral structures of domains II and VI are required for the conformational changes happening between the two catalytic steps (Chanfreau *et al.*, 1996).

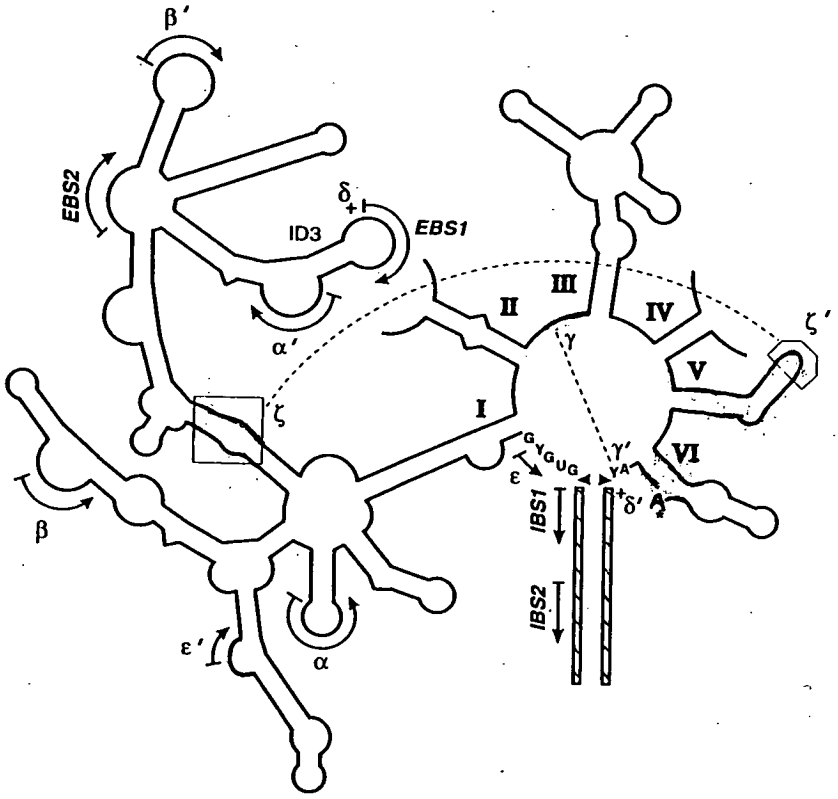
***Similarities with other groups of introns:*** several analogies between group II introns and nuclear pre-mRNA introns exist. The hairpin VI of group II introns is similar to a double-stranded structure located in the yeast spliceosome and constituted by the base-pairing between the pre-mRNA branchsite sequence UACUAAC and the U2 snRNA. Also, in both structures the branchpoint adenosine is bulged out from the intron. The ε-ε' interaction in group II introns is similar to that between the 5' splice site of the nuclear pre-mRNA introns and the U6 snRNA (Jacquier and Jacquesson-Breuleux, 1991). The group II intron catalytic domain V is similar to the U2 snRNA-U6 snRNA interaction in the spliceosome (Jarrell *et al.*, 1988). The terminal loop of group II intron domain I (subdomain ID3) contains the exon binding site 1 (EBS1) which is complementary to the 3' end of the 5' exon (the intron binding site 1, IBS1). In many group II introns, this subdomain ID3 can pair with bases at the 5' end of the 3' exon, contributing to the 3' splice site selection and activation (Jacquier and Jacquesson-

**Figure 2. Group II introns.**

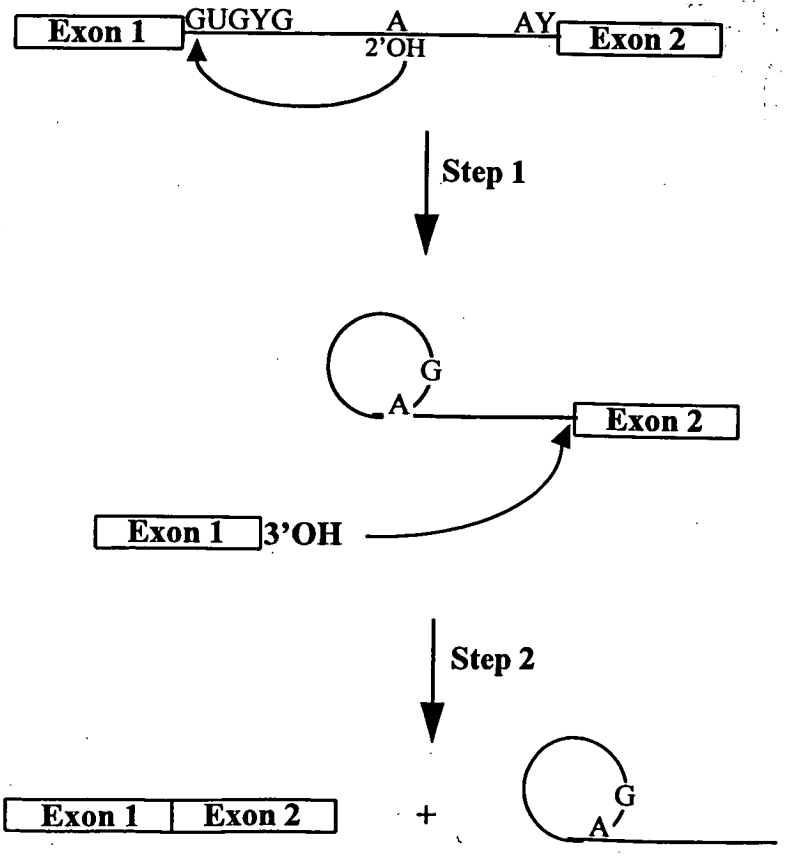
**A:** Secondary structure of intron *ai5γ* from the *S. cerevisiae* mitochondrial genome (Michel *et al.*, 1995). The intron is indicated by the plain line and the exons by the rectangles. Tertiary interactions are indicated by dashed lines, curved arrows, and/or greek lettering. The consensus sequence of the 5' splice site is GUGYG; the branchpoint is A; the 3' splice site is AY. Most group II introns have a bulging A (branchpoint, indicated on this diagram with an asterisk) on the 3' side of the basal helix of domain VI at either seven or eight nucleotides from the 3' splice site. Ribozyme domain I includes sequences called the exon-binding sites (EBS1 and EBS2) and base-pairs to the 5' exon sequences called intron-binding sites (IBS1 and IBS2 respectively). Integrity of the EBS1-IBS1 pairing is essential both for lariat formation and stability of the intermediate complex. Also EBS-IBS interactions are involved in aligning the splice sites. Domain V is the most highly conserved and is believed to interact with domain I to form the catalytic core (Koch *et al.* 1992). All of the group II intron ORFs are inserted in intron domain IV. Domain IV is highly variable and often quite short. Relatively little is known of the exact role of the remaining domains II and III. They are presumed to contribute to the folding of the ribozyme or to protein binding. Y: pyrimidine; ID3: internal domain 3.

**B:** Schematic representation of the group II intron splicing pathway. The first step involves the formation of a intron-lariat-exon 2 intermediate in which the nucleotide at the 5' end of the intron is covalently linked (via a 2'-5' phosphodiester bond) to a nucleotide (generally an A-residue) located at the branchpoint site. The second step is characterised by exon ligation coupled to cleavage at the 3' splice site and release of the intron-lariat.

A


β′
EBS2
δ+1
ID3
EBS1
α′
ζ
I
II
III
IV
V
VI
ζ′
γ
γ′
ε
GYGUG
YA
δ′
IBS1
IBS2
β
ε′
α

B


Exon 1
GUGYG
A
2'OH
AY
Exon 2
Step 1
G
A
Exon 2
Exon 1
3'OH
Step 2
Exon 1
Exon 2
+
G
A

Breuleux, 1991; Sontheimer and Steitz, 1993). This subdomain functionally resembles the conserved single-stranded loop of the U5 snRNA which interacts with exon sequences at both the 5' and the 3' splice sites (Newman and Norman, 1992; Steitz, 1992). In nearly all group II introns, contacts between IBS and EBS are established as Watson-Crick base-pairing interactions (Michel *et al.*, 1989b), whereas the highly conserved loop I of the U5 snRNA does not display significant exon sequence specificity and can be promiscuous in its base-pairing with exon sequences (Newman and Norman, 1992). Indeed, the U5 snRNA can bind to different exon sequences even with drastically altered loop I sequences (O'Keefe *et al.*, 1996, 1998). Hetzer *et al.* (1997) have trans-activated the autocatalytic group II intron *bl1* (yeast mitochondria group II intron) lacking the subdomain ID3 with the U5 snRNA loop I, suggesting that ID3 and the U5 snRNA conserved loop are functional counterparts. Some authors have proposed that these numerous resemblances between group II introns and nuclear pre-mRNA introns reflect a common origin. According to an evolutionary hypothesis, RNA domains or subdomains of self-splicing group II introns would have been successively detached from the catalytic core to become separately encoded (Cech, 1990; Michel, 1989; Nilsen, 1994). Despite these similarities, the spliceosomal U4 and U1 snRNAs are left with no plausible counterparts. The last nucleotide of group II intron forms a classical base-pairing with one of the residues in the segment connecting domains II and III (the γ-γ' pairing, Figure 2A). No such interaction has been uncovered in spliceosomal introns, whose last intron nucleotide instead seems to interact with the first by nonclassical base-pairing during the second step of splicing (Parker and Siliciano, 1993). Also, group II introns have 5' and 3' consensus sequences (GUGYG and AY, respectively) which resemble those found in nuclear pre-mRNA introns, but unlike them, they have a complex conserved structure required for splicing. These differences between nuclear pre-mRNA and group II introns are often interpreted as evidence that the two systems have diverged to such an extent that tertiary structure became somehow altered.

## 1.3 Group III Introns

In 1989, Christopher and Hallick discovered the first group III intron in the *Euglena gracilis* chloroplast operon encoding for ribosomal protein L5. Since then, several group III introns have been discovered in the *Euglena gracilis* genome and in the plastid of the heterotrophic flagellate *Astasia longa* (Siemeister *et al.*, 1990).

***Common features:*** group III introns are small and uniform in size (~100 nt), except for the 1503 nt long internal group III intron of *psbC* twintron. Group III introns are abbreviated group II introns. They lack domains II-IV, but have retained group II-like 5' and 3' boundary sequences, and at their 3' end a domain VI-like structure with a bulged A used during lariat formation. Some group III introns also contain the subdomain ID3 (Copertino and Hallick, 1993). The mechanism of splicing of group III introns resembles that of group II; both splice via a lariat intermediate containing an internal 2'-5' phosphodiester bond (Parker *et al.*, 1987; Jacquier, 1990; Copertino *et al.*, 1992).

***Similarities with other groups of introns:*** the U residue in the second position and the G residue in the fifth position of the 5' splice site boundaries of group III introns (5'-NUNNG), group II introns (5'-GUGYG) and nuclear pre-mRNA introns (5'-GUNNG) are conserved (N being any nucleotide). Based on the striking similarities between group II and group III introns, some authors have proposed that many of the *cis* elements necessary for group II intron excision, and lacking in group III introns, are supplied in *trans*. For example, the group III intron of the *psaA* mRNA of the chlamydomonas chloroplast can be activated when a domain V from a group II intron is supplied in *trans* (Goldschmidt-Clermont, 1991). This suggests an evolutionary relationship between group II and group III introns (Jarrell *et al.*, 1988; Goldschmidt-Clermont, 1991; Hong and Hallick, 1994). Also the requirement for *trans*-acting RNAs for group III intron splicing would be an obvious parallel to the role of snRNAs in nuclear pre-mRNA splicing.

## 1.4 Nuclear Pre-mRNA Introns and Their Splicing Mechanism

In the nucleus of higher eukaryotic cells, many protein-encoding genes are interrupted by at least one or more (up to 65) introns of the major class. Many human diseases are caused by mutations that interfere with nuclear pre-mRNA splicing. Thus, conserving information for accurate splicing is a constraint on genetic systems. On the other hand, the presence of multiple introns allows the choice of several combinations of exons by alternative splicing, producing from an identical pre-mRNA different mature RNAs which extends the genome information capacity.

The yeast *Saccharomyces cerevisiae* (referred to yeast in this work), differs from higher eukaryotes in that split genes seem to be an exception. Only ~3% of the yeast nuclear genes contain an intron, whereas in the fission yeast *Schizosaccharomyces pombe* at least half of the genes are interrupted. In the former, most of the intron-containing genes possess one intervening sequence, with the expection of the *MATa1* and the *RPL8A* genes which contain two small introns. In *S. cerevisiae,* nuclear introns are generally small, the largest being 1001 bases long in the *DBP2* gene (Iggo *et al.*, 1991).

### 1.4.1 Description of the nuclear pre-mRNA introns excised by the U2-type spliceosome

Introns are defined by the 5' and 3' splice sites and the branchpoint sequence elements which are shown in Figure 3.

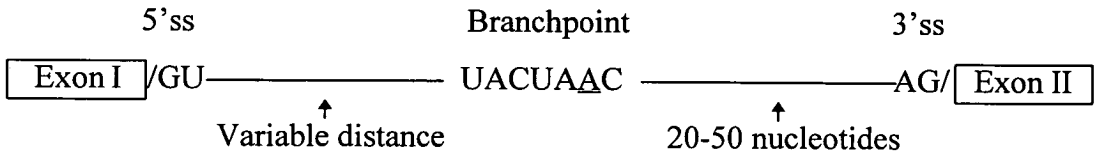
The importance of the consensus elements has been demonstrated *in vitro* and *in vivo* through site-directed mutagenesis and by analysing the effects of naturally occurring mutations. Mutations within the 5' splice site generally abolish splicing in yeast (Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986), or lead to the use of proximal (cryptic) sites in mammalian systems (Aebi, 1986, 1987). In yeast mutations within the branchpoint sequence generally abolish splicing, however some mutations allow the

**Figure 3. Schematic representation of a nuclear pre-mRNA intron indicating the positions and the sequence of the three conserved elements.**

**A:** 5' splice site (5' ss), branchpoint where A is the site of branch formation and 3' splice site (3' ss) are indicated on this diagram. The distance between the 5' splice site and the branchpoint varies. However, splicing efficiency is affected if this distance is shorter than 45 nucleotides or greater than 700-1000 nucleotides.

**B:** Splice site consensus sequences for yeast and higher eukaryotes introns are indicated. In higher eukaryotes, two subclasses (GT-AG and AT-AC) of nuclear pre-mRNA introns are excised by the U2-spliceosome, however the GT-AG introns are the most frequently found in the genome (for review see Sharp and Burge, 1997).
R: purine; Y: pyrimidine, /: splice site.

**A**




**B**

| Organisms | 5' ss consensus sequence | 3' ss consensus sequence |
|---|---|---|
| **Yeast** | /GUAUGU | CAG/ |
| **Mammals** | AG/GURAGU | YAG/G |
| **Plants, Drosophila and vertebrates** | ARG/AUAAGU | YAC/ |

first step to occur but block the second step of splicing (Fouser and Friesen, 1986; Jacquier and Rosbach, 1986).

Additional *cis*-acting elements that influence splicing include exon sequences flanking the intron (Newman and Norman, 1991, 1992). In yeast a polypyimidine tract located between the branchpoint and the 3' splice site, though not strictly requisite, greatly enhances the efficient usage of an adjacent AG dinucleotide in alternative 3' splice site competition assays (Patterson and Guthrie, 1991).

Unlike yeast, mammalian introns have defined but less well conserved elements. These differences are more likely to reflect the differences between the splicing machineries. Indeed, transcripts of genes from other organisms are generally not accurately spliced in *S. cerevisiae* because the intron sequences are not appropriate (Beggs *et al.*, 1980), whereas yeast transcripts are spliced *in vitro* in mammalian systems although alternative branchpoint sequences are used.

### 1.4.2 The basic nuclear pre-mRNA splicing reaction

This reaction is shown in Figure 4. It involves two trans-esterification steps with the formation of a lariat-intermediate (reviewed in Moore *et al.*, 1993). Although these reactions do not themselves seem to need any exogenous source of energy, ATP hydrolysis is required for nuclear pre-mRNA splicing. Spliceosome formation and/or some spliceosome rearrangements are more likely to be responsible for this ATP-dependency. GTP-hydrolysis is required for the function of the U5 snRNP specific protein U5-116 kD/Snu114p indicating that GTP may be needed for splicing (Fabrizio *et al.*, 1997).

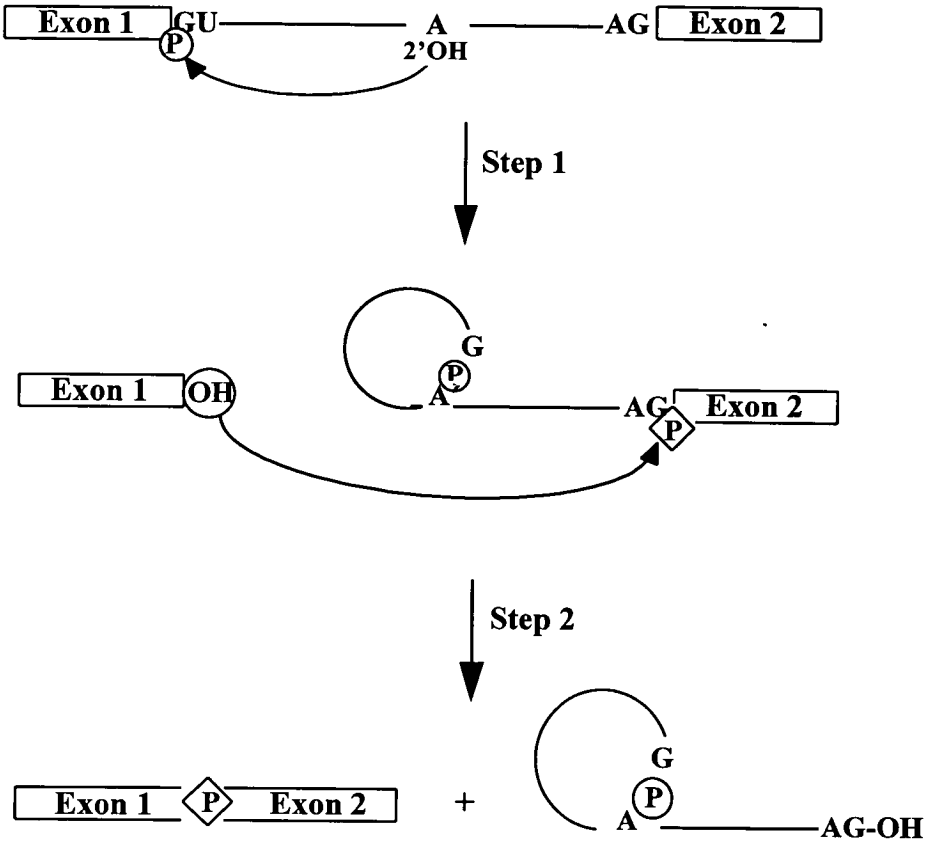



**Figure 4. Nuclear pre-mRNA splicing reaction.**
Exon sequences (rectangles), intron sequences (lines) and conserved nucleotides and the branchpoint as well as reactive phosphates and hydroxyl groups are indicated. In step 1, the 5' splice site phosphodiester bond is cleaved by nucleophilic attack of the 2' hydroxyl group of the branchpoint A residue on the phosphate bond at the 5' splice site. This yields a 5' exon intermediate with a free 3'-OH terminus (Exon 1-OH). Concurrently, the 5' end of the intron is covalently linked via a 2'-5' phosphodiester bond to the bulged adenosine of the branchpoint sequence, this forms the so-called lariat-intron-exon 2 intermediate. In step 2, cleavage at the 3' splice site occurs as a result of a second nucleophilic attack of the free hydroxyl of exon1 on the phosphodiester bond. Consequently, the intron-exon 2 lariat-intermediate is resolved and the exons ligated together through a 3'-5' phosphodiester bond. This yields two products: the spliced exons (or mature RNA) and an excised lariat-intron product that is rapidly degraded *in vivo*.

### 1.4.3 Small nuclear ribonucleoprotein particles of the U2-type splicing machinery

Unlike group I and group II introns, conserved sequences within the nuclear pre-mRNA are not sufficient to catalyse the accurate and specific removal of introns. The paucity of highly conserved structural elements in nuclear introns appears to be compensated by the involvement of a large number of *trans*-acting factors. These factors interact in a highly ordered fashion with each other and/or on the nuclear pre-mRNA intron to form a large ribonucleoprotein complex called the spliceosome. The spliceosome is constituted by a set number of small nuclear ribonucleoprotein (snRNP) particles and associated splicing factors (proteins which interact transiently or stably with the spliceosome). SnRNP particles are constituted by one or two small nuclear RNA (snRNA) that are associated with proteins.

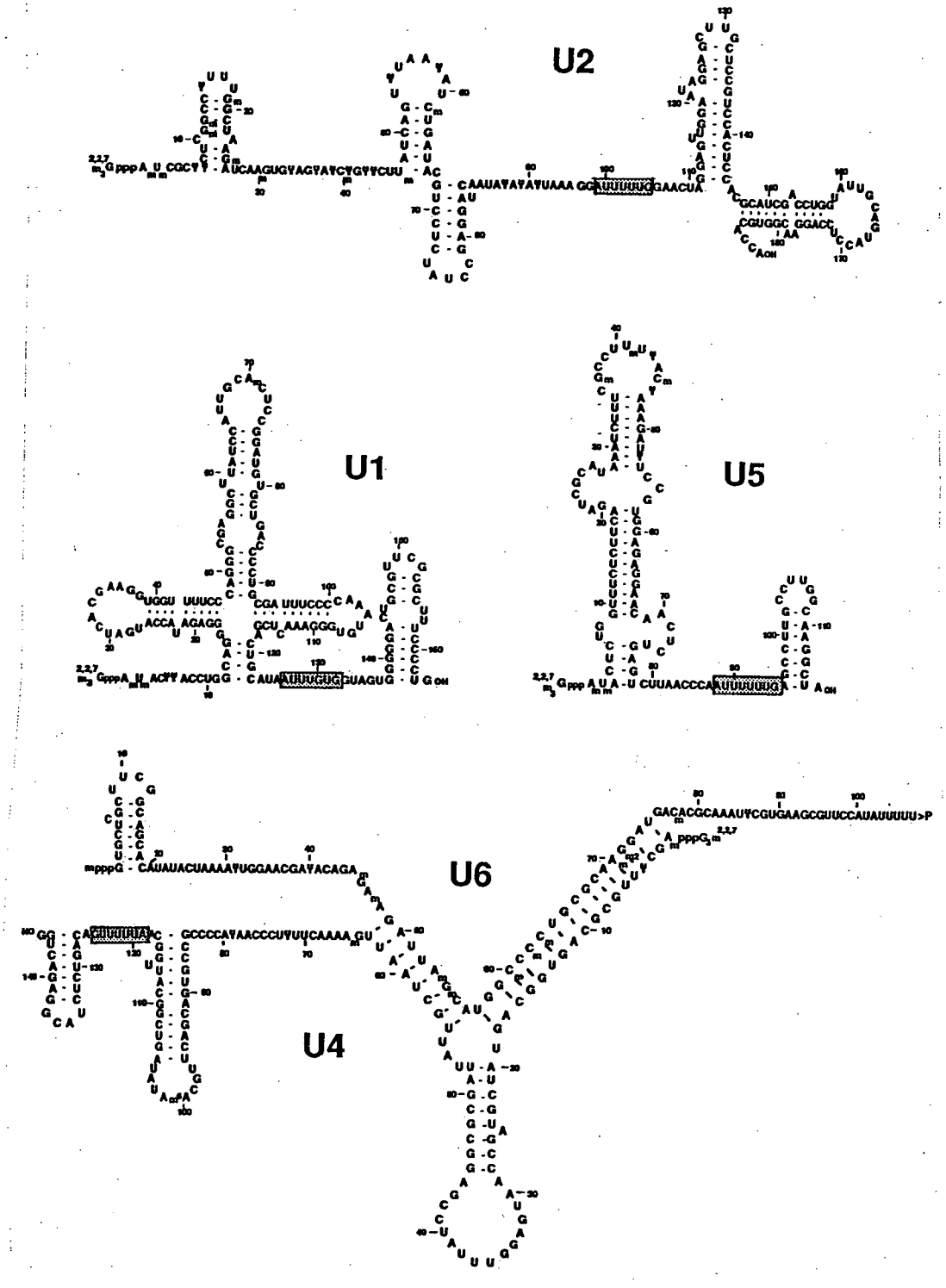
#### 1.4.3.1 The snRNAs involved in nuclear pre-mRNA splicing

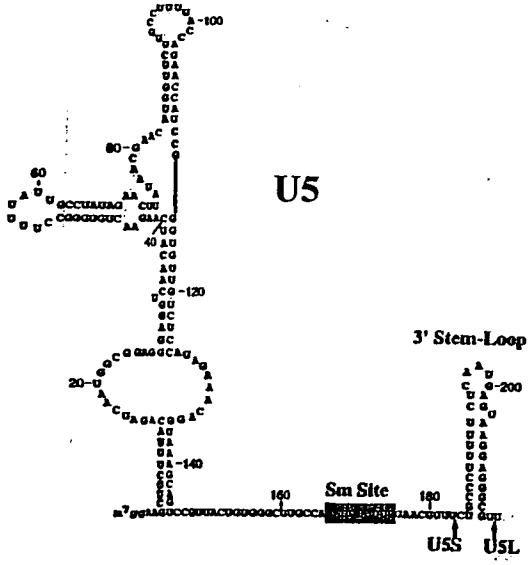
***Description of the snRNAs***

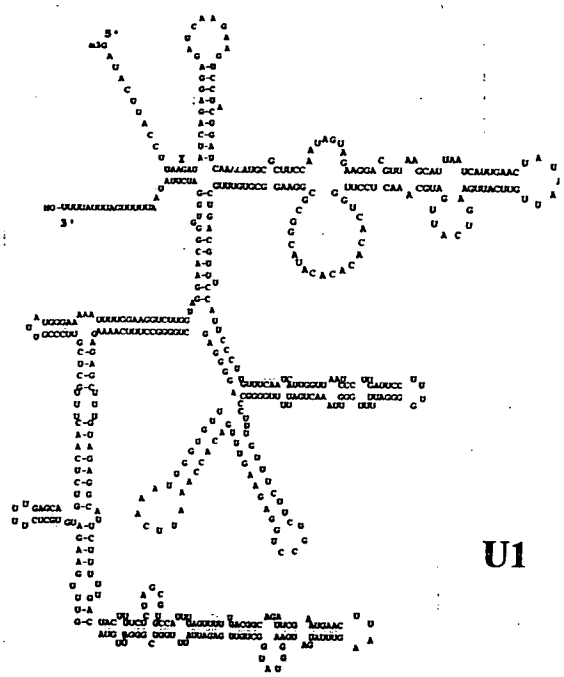
Five spliceosomal snRNAs named U1, U2, U4, U5 and U6 snRNAs, are involved in the splicing of the major class of introns. These are the most abundant of the snRNAs. Although their primary sequence varies between species, secondary structures are highly conserved. The secondary structure of the mammalian and yeast snRNAs are shown in Figures 5 and 6 respectively.

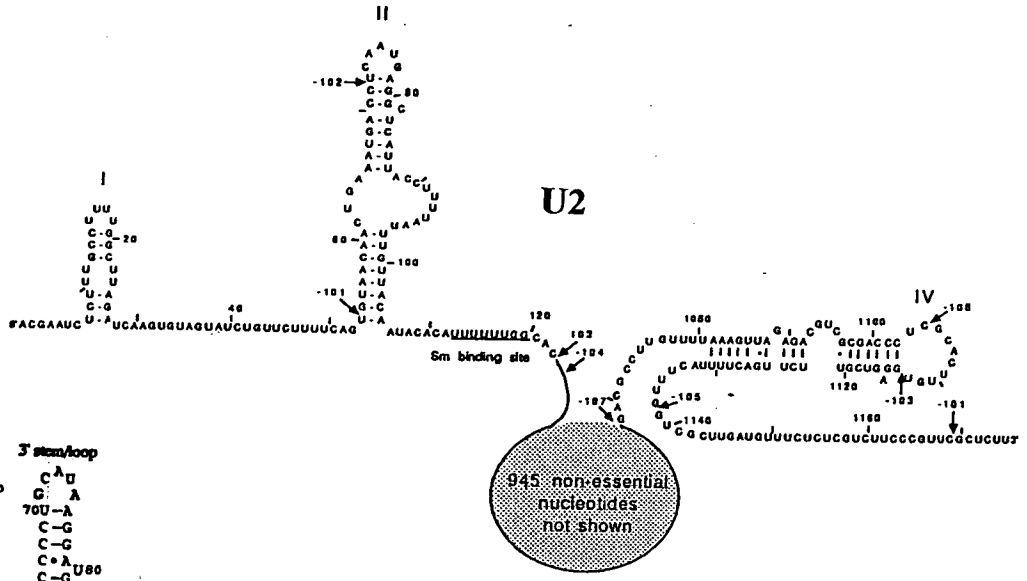
With the exception of U4 and U6 snRNAs which are associated by extensive base pairing interactions and packaged into a single particle (the U4/U6 snRNP), only one snRNA molecule is present per snRNP.

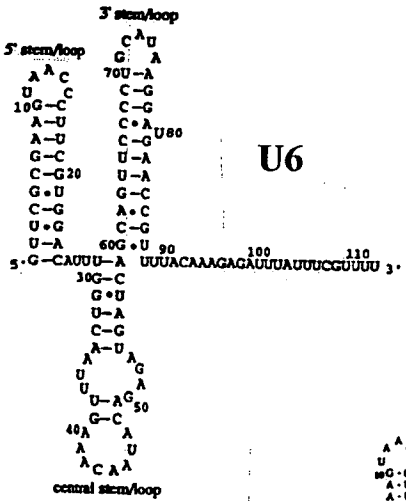
The U1, U2, U4 and U5 snRNAs are transcribed by the RNA polymerase II and they possess at their 5' end a trimethylated cap structure (2, 2, 7-trimethylguanosine; Reddy and Busch, 1988). The U6 snRNA differs from the other snRNAs in that it is

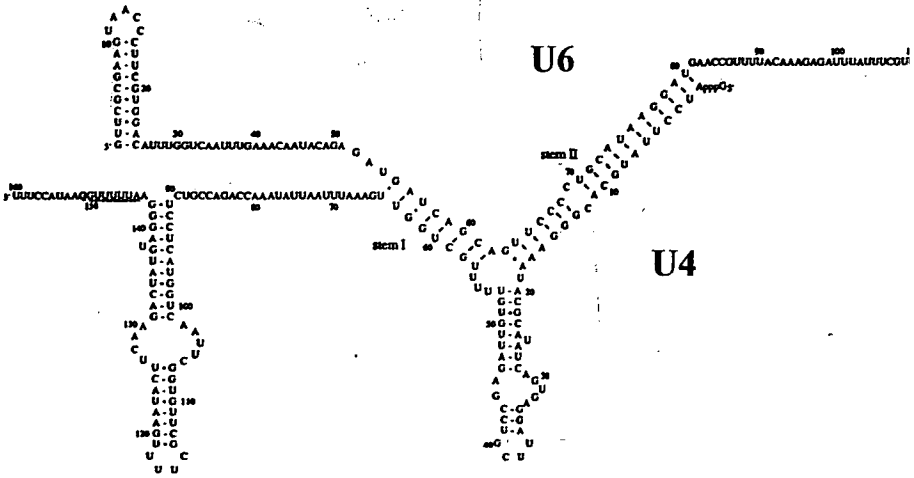

U2
U1
U5
U6
U4


U5
3' Stem-Loop
Sm Site
U5S
U5L


U1


U2
I
II
IV
Sm binding site
945 non-essential nucleotides not shown


U6
3' stem/loop
5' stem/loop
central stem/loop


U6
U4
stem I
stem II

transcribed by the RNA polymerase III and consequently lacks the trimethylated cap structure. Instead, it possesses at its 5' end a γ-methyl phosphate cap.

*The U1 snRNA*. It is characterised by a 5' single-stranded region containing a sequence complementary to the 5' splice site of consensus sequence pre-mRNAs. This region is followed by a long-range interaction of 4-5 base-pairs that closes off a central domain comprising three contiguous stem-loop structures designated I, II and III (Guthrie and Patterson, 1988). The 3' terminal domain consists of the single-stranded Sm binding site (see this section for definition), followed by a terminal stem-loop structure, designated IV. The length of each domain is generally conserved between organisms, the total size of U1 snRNA varying only between 165 nt (mouse) and 147 nt (*Schizosaccharomyces pombe*). The yeast U1 snRNA is slightly different since it is 568 nt in length and lacks the 3' terminal stem-loop element.

*The U2 snRNA*. Its size greatly varies among the different organisms studied and ranges from 148 nt (trypanosomes) to 1175 nt (*S. cerevisiae*) (Guthrie and Patterson, 1988). The U2 snRNA can be divided into four stem-loop domains (I-IV), with single-stranded regions separating stem I from stem II and stem II from stem III. While the third stem-loop is highly conserved in length (~10 base-pairs) and loop sequence (CUUG), it is completely missing in yeast where it is replaced by an insertion of ~1 kb (Shuster and Guthrie, 1988). Surprisingly, this insertion can be deleted without any impact on yeast viability. Despite the large size of the yeast U2 snRNA, it can be folded into a structure that resembles the structure originally proposed for the metazoan U2 snRNA by Keller and Noon (1985). The branchpoint recognition sequence identified by Parker *et al.* (1987) is located in the single-stranded region immediately downstream of stem-loop I. Within this single-stranded region, nucleotides 28-47 are absolutely conserved (Guthrie and Patterson, 1988). The Sm binding site is located downstream from stem-loop II.

*The U6 snRNA*. It is the most highly conserved of the spliceosomal snRNAs, and has been postulated to have a fundamental role in nuclear pre-mRNA splicing (Das *et al.*, 1987; Brow and Guthrie, 1988). The 5' terminal domain (nucleotides 1-27) contains a hairpin structure highly variable in length among species, which is responsible for essentially all the differences in size of the U6 snRNAs. The contiguous stretch of conserved sequences (nucleotides 55-80) in U6 snRNA has been proposed to base-pair with the U4 snRNA to form two contiguous stems (Bringmann *et al.*, 1984; Hashimoto and Steitz, 1984; Rinke *et al.*, 1985; Das *et al.*, 1987; Siliciano *et al.*, 1987; Kiss *et al.*, 1988). The 3' terminal domain (nucleotides 81-112) is divergent in sequence but constant in length. The sequencing of the U6 snRNA revealed that the 3' end was heterogeneous and contains 4, 5, 6, or 7 contiguous U residues (Brow and Guthrie, 1988). In yeast, point mutations (Fabrizio and Abelson, 1990; Madhani *et al.*, 1990, Madhani and Guthrie, 1992) or thiophosphate substitutions (Fabrizio and Abelson, 1992) in the two invariant sequences (-$A_{47}$CAGAG$A_{53}$- and -$A_{59}$G$C_{61}$-) block either the first or the second step of the splicing reaction. The same effect was observed for mutations in the corresponding sequences in *Xenopus* U6 snRNA (Vankan *et al.*, 1992). These two sequences were less sensitive to mutation in human cells, reflecting differences in the requirement of U6 snRNA sequences between human and yeast despite their overall similarity (Datta and Weiner, 1993).

*The U4/U6 snRNAs*. The U4 snRNA molecule has been proposed to contain six domains: stem II, 5' stem-loop, stem I, central region, 3' stem-loop and 3' terminal region. The 3' terminal domain of the U4 snRNA contains the Sm binding site, which is followed in all organisms except in *S. cerevisiae* and *K. lactis* by a 3' terminal stem. The most remarkable aspect of the U4 snRNA/U6 snRNA is the interaction domain called domain Y (Brow and Guthrie, 1988; Guthrie and Patterson, 1988). Analysis of RNA content in yeast cells has shown that U6 snRNA is present at four- to five-fold molar excess relative to U4 snRNA (Li and Brow, 1993), while, 95% of the U4 snRNA is complexed to U6 snRNA (Fortner *et al.*, 1994). The U6 snRNA sequence bound to the U4 snRNA shows higher conservation than does the comparable sequence of the U4

snRNA. U4 snRNA probably sequesters a catalytically active domain of U6 snRNA until other conditions of spliceosome assembly have been satisfied. The U4 snRNA could function as a negative effector (Yean and Lin, 1991; Moore *et al.*, 1993). It has been suggested that the region of U6 snRNA interacting with the U4 snRNA may be involved in catalysis (Brow and Guthrie, 1988). This hypothesis was supported by mutagenesis studies and the discovery of an intron within the gene encoding for the U6 snRNA in the fission yeast *Schizosaccharomyces pombe*. This intron was located in the middle of the most highly conserved region of U6 snRNA (Tani et Ohshima, 1989), which might reflect the proximity of that region to the catalytic centre of the spliceosome.

*The U5 snRNA*. It varies in size over a two-fold range (from 107 nt in dinoflagellates to 214 nt in *S. cerevisiae*). The U5 snRNA appears to be the most highly structured of the spliceosomal snRNAs. The molecule appears to be organised into two domains: a 5' terminal compound helix and a 3' terminal domain containing a single-stranded region which includes the Sm binding site and a stem-loop. The long helical element located at the 5' extremity is designated stem I. This stem is divided into several substructures with the maximum conservation found in the top portion of the terminal loop I (among the 11 nt constituting this loop, 9 are invariant in sequence). The 3' terminal stem-loop of U5 snRNA varies considerably in size and sequence in different organisms. In yeast, half of the U5 snRNA transcripts terminate just prior to the 3' terminal stem-loop producing two forms of U5 snRNA which are called U5S (for short) and U5L (for long) (Patterson and Guthrie, 1987; Frank *et al.*, 1994). Both forms are incorporated into snRNP particles (Madhani *et al.*, 1990), yet the U5L is dispensable *in vivo* (Haltiner-Jones and Guthrie, 1990; Frank *et al.*, 1994). Mammalian U5 snRNA is extremely resistant to micrococcal nuclease and RNase H cleavage, suggesting that even the single-stranded RNA sequences are largely inaccessible.

### *Processing of the snRNAs*

During snRNP biogenesis, snRNAs undergo several changes which include: transcription as precursor snRNAs containing at their 3' end an extension of 1-16 nt (Kleinschmidt and Pederson, 1987; Madore *et al.*, 1984a, 1984b), assembly with the common proteins, trimming at the 3' end of the snRNA; extensive internal modifications; 5' end capping (Dahlberg *et al.*, 1988; Mattaj, 1988). The processing of snRNAs is very complex and involves several enzymatic activities which are not yet fully understood. Only a brief overview of these processing steps is given in this section.

*Formation of pseudouridine*. SnRNAs are highly modified post-transcriptionally (methylation of bases, formation of pseudouridine). Pseudouridine (Ψ) formation represents by far the most abundant of these modifications. In HeLa cell extracts, the formation of Ψ involves multiple Ψ synthase activities that specifically recognise the U1, U2, and the U5 snRNAs (Patton, 1991, 1993, 1994). Efficient Ψ formation in U5 snRNA requires the binding of the Sm proteins (Patton, 1991), whereas Sm proteins are not needed for the formation of Ψ in U2 snRNA (Kleinschmidt *et al.*, 1989; Patton *et al.*, 1994). U4 and U6 snRNAs are also modified. Intriguingly, the formation of Ψ in the U6 snRNA depends upon its interaction with the U4 snRNA (Zerby and Patton, 1996). The function of Ψ in spliceosomal snRNAs is unknown, however it is noteworthy that Ψ are found in regions of snRNAs known to be necessary for the function of snRNPs in splicing (Steitz *et al.*, 1988; Guthrie and Patterson, 1988). The requirement for Ψ modification for the *in vitro* function of the snRNA varies. In HeLa cell extracts, modifications are required for the activity of the U2 snRNA while they are dispensable for the activity of the U5 snRNA (Ségault *et al.*, 1995). This difference in the requirement of Ψ may be due to the fact that U2 snRNA contains a larger percentage of Ψ; therefore the absence of Ψ might have a greater impact on its function (Ségault *et al.*, 1995).

*Trimming of the 3' end.* Transcription of snRNAs produces precursor molecules which contain at their 3' end an extension of 1-16 nts (Kleinschmidt and Pederson, 1987; Madore *et al.*, 1984a, 1984b). These longer forms are shortened by mechanism(s) yet unknown. Most of this extension is removed in the cytoplasm where Sm protein binding and cap hypermethylation of the snRNAs occur concomitantly (Madore *et al.*, 1984a; Kleinschmidt and Pederson, 1987; Neuman de Vegvar and Dahlberg, 1990). In vertebrates, proper snRNA 3' end formation requires transcription from snRNA promoters only, since the transcription of snRNAs from a mRNA promoter results in aberrant processing of the transcript. Thus the production of a precursor snRNA with an extended 3' tail seems linked to transcription (Ciliberto *et al.*, 1986; Hernandez and Weiner, 1986; Neuman de Vegvar *et al.*, 1986). Also, the 3' end maturation of snRNAs in metazoans is constrained to occur within a short distance from the promoter (Ramamurthy *et al.*, 1996). In contrast, in yeast the 3' end shortening is not affected by substitution of the snRNA promoter with elements of mRNA promoters. Also, there is no restriction in the distance between the promoter and the 3' end of the snRNA. These differences probably reflect the different 3' end snRNA processing pathway(s) in metazoans and in yeast. The study of the 3' end processing of the yeast U5 snRNA indicates that the formation of the U5S and U5L involves two distinct pathways with the RNase III being required for the production of the intermediates (Chanfreau *et al.*, 1997). The complete mechanisms involved in the generation of U5L and U5S has not yet been deciphered. Only an exonuclease is suspected to be needed for the generation of U5L (Chanfreau *et al.*, 1997).

*The Sm site and the association of the Sm proteins.* U1, U2, U4 and U5 snRNAs share a common structural motif called the Sm site. This motif serves as the primary binding site for the common (also called Sm or common or core) snRNP proteins (see section 5.3.2 for description; for review see Mattaj, 1985, 1988; Hamm *et al.*, 1990). This site has been identified as a single-stranded U-rich sequence ($RAU_{3-6}GR$; R: purine base) which is most often flanked by two hairpin structures (Branlant *et al.*, 1982). Saturation mutagenesis demonstrates that only three of the ten consensus

nucleotides of the Sm binding site are critical (UAUUUUUUGG; crucial nucleotides are underlined), and their mutation is associated with a conditional or lethal phenotype (Jones *et al.*, 1990; Hu *et al.*, 1995). The nucleotides bordering the critical residues are relatively insensitive to site-specific mutation, whereas multiple mutations are associated with severe growth defects. Complete deletion or substitution of the yeast U4 or U5 snRNA Sm site is lethal, confirming the functional importance of this sequence. However, unexpectedly, the deletion of the Sm site in U4 snRNA from HeLa cells has little effect on *in vitro* splicing or spliceosome assembly, but abolishes Sm protein binding (Wersig *et al.*, 1992).

The association of the common snRNP proteins depends on this Sm site and the presence of only one flanking stem-loop (which may be positioned either on the 5' or 3' side of the Sm site). The requirement of only one hairpin structure is illustrated by the fact that one of the forms of the yeast U5 snRNA (U5S) terminates shortly after the Sm site and thus lacks a 3' flanking hairpin (Guthrie *et al.*, 1988). Although these structures are identical, they are not interchangeable among snRNAs. They might finely tune protein-protein interactions between the Sm proteins themselves (Jarmolowski *et al.*, 1993). The binding of the Sm proteins to the snRNAs is required for the hypermethylation of the 5'-5' 7-methylguanosine (7-mG) cap structure of the snRNAs, producing a hypermethylated 2,2,7-trimethylguanosine cap. The core proteins probably provide a binding site for the *trans*-active snRNA-(guanosine-*N2*)-methyltransferase (Plessel *et al.*, 1994).

Proper Sm protein assembly, cap hypermethylation and 3' end processing are important for subsequent nuclear import of the particles (reviewed by Mattaj, 1988; Zieve and Sauterer, 1990; Izaurralde and Adam, 1998). Just before and/or after nuclear import, snRNP-specific proteins associate with the individual snRNP precursors to complete biogenesis. Whether the Sm proteins have a function in nuclear pre-mRNA splicing is still unknown.

Very little is known about the biogenesis of the U6 snRNP particle and the processing of the U6 snRNA. In HeLa cells and in *Xenopus* oocytes the U6 snRNA does not leave the nucleus(Boelens *et al.*, 1995; Pante *et al*, 1997), whereas in L929 mouse fibroblasts where it has been detected in the cytoplasm as well as in the nucleus (Zieve *et al.*, 1988; Fury and Zieve, 1996). Canonical Sm proteins do not bind to free U6 snRNA (Lührmann *et al.*, 1990). However U6 snRNP-specific proteins (see section 5.3.2) which share homology with the Sm motif of the core proteins have been identified and some have been characterised (Cooper *et al.*, 1995; Séraphin, 1995; Racine-Fromont *et al.*, 1997).

### 1.4.3.2 The protein composition of the snRNPs involved in nuclear pre-mRNA splicing

Two types of proteins associated with the spliceosomal snRNAs are distinguished: the common (also called the core or Sm proteins) found in all snRNP particles (with the exception of the U6 snRNP), and the snRNP particle-specific proteins which are exclusively associated with a particular snRNA or a complex of snRNAs.

#### *The Sm proteins*

They are referred to as Sm proteins because they possess an antigenic determinant recognised by auto-antibodies isolated from patients (so-called anti-Sm sera after the patient Smith) suffering from the autoimmune disease systemic lupus erythematosus. Anti-Sm sera immunoprecipitate also yeast Sm proteins showing that immunological determinants of the core proteins are conserved (Siliciano *et al.*, 1987; Tollervey *et al.*, 1987). The reconstitution *in vitro* of mammalian snRNP particles with purified Sm proteins shows that they are essential for splicing. Also, Sm proteins from the U2 snRNP or the U1 snRNP particle can reconstitute a functional U5 snRNP

lacking its own Sm proteins, demonstrating that common proteins are functionally interchangeable (Ségault *et al.*, 1995).

A list of all the mammalian and yeast Sm proteins is presented in Table 1.

**Table 1. List of the mammalian and yeast Sm (common/core) proteins.**

| *Common SnRNP Particle Proteins* | | |
|---|---|---|
| | **Human** | **Yeast** |
| **Common name** | **Size** | **Size - other name(s)** |
| SmB' | 29 kD | |
| SmB | 28 kD | 22 kD - YER029 |
| SmD1 | 16 kD | 16 kD - YGR074 |
| SmD2 | 15.5 kD | 12 kD - YLR275 |
| SmD3 | 18.5 kD | 11 kD - YLR147 |
| SmE | 12 kD | 10 kD - YOR159 |
| SmF | 11 kD | 9.7 kD - YPR182-SMX3 |
| SmG | 9 kD | 8.5 kD - YFL018-SNP2-SMX2 |

The different names given for the yeast proteins are indicated. The corresponding human and yeast homologues are indicated by a dashed line.

Two sequence motifs, denoted Sm motif 1 and 2, are found in all of the known Sm proteins from organisms as diverse as yeast, nematodes, insects, plants and higher vertebrates (Cooper *et al.*, 1995; Hermann *et al.*, 1995; Séraphin, 1995). These Sm motifs are important for Sm protein-protein interactions (Hermann *et al.*, 1995).

*The mammalian Sm proteins*. With the exception of F (pI: 4.6), all the core proteins are basic with a pI value between 8 and 10. B and B' are encoded by the same gene (Van Dam, 1989; Chu and Elkon, 1991; for review see Will *et al.*, 1993). Another polypeptide named N, is also shared by the snRNP particles but expressed specifically in neural tissues (McAllister *et al.*, 1989).

Sm proteins bind to the Sm site of the U1, U2, U4 and U5 snRNAs, and to one another, to form a highly stable snRNP core structure. In the snRNP particles, the core proteins protect a 15-25 nt long region of the Sm site against hydrolysis by micrococcal nuclease (Liautard *et al.*, 1982). To date the only known direct interaction between a core protein and a snRNA is that demonstrated by UV-crosslinking analysis between the human U1 snRNA Sm site (more precisely the AAU sequence of the AAUUUGUGG Sm motif) and the SmG protein (Heinrichs *et al.*, 1992). Protein-protein, rather than protein-RNA interactions appear to dominate in determining the structure of the snRNP core.

Some Sm proteins are able to interact together in the absence of snRNA: D1.D2 (Lehmeier *et al.*, 1994); B.D3 and B'.D3 (Hermann *et al.*, 1995); E.F, E.G, E.F.G and D1.D2.E.F.G ( Fisher *et al.*, 1985; Sauterer *et al.*, 1990; Raker *et al.*, 1996; Plessel *et al.*, 1997). The assembly of the Sm proteins to form the core structure has been well studied in mammalian *in vitro* systems (Fisher *et al.*, 1985; Sauterer *et al.*, 1990; Raker *et al.*, 1996; Plessel *et al.*, 1997). The core proteins E, F, G interact together to form a (E.F.G) complex. Given the sedimentation coefficient and observation of the core structure by electron microscopy, it is more likely that two (E.F.G) complexes are present per snRNP particle (Raker *et al.*, 1996; Plessel *et al.*, 1997). This multimer probably provides a structural platform for snRNP core assembly. The ultrastructural similarity of U1, U2, U4 and U5 snRNP cores indicates the presence of the (E.F.G) dimer in each of these particles (Plessel *et al.*, 1997). D1 and D2 interact together and bind to the hexamer (E.F.G)2. This heteromeric complex can associate with a monomethylated U1 snRNA forming the subcore snRNP. The subcore formation is not restricted to the U1 snRNA, and is detected with U2, U4 and U5 snRNAs (Raker *et al.*, 1996). Then B, B' and D3 associate with the subcore to form the core particle corresponding to the U snRNA and the eight common proteins (Hermann *et al.*, 1995). It is not known if both B and B'or only one of these proteins assemble *in vivo* into the subcore particle. B, B' and D3 are required for the conversion of the U1 snRNA's $m^7G$ cap to its hypermethylated $m_3G$ form. The stochiometry of D1.D2 and B/B'.D3 in the snRNP particles remains unknown.

*The yeast Sm proteins*. The much lower abundance of snRNP particles in yeast compared to metazoans has made biochemical purification of yeast core proteins a difficult task. Also, their identification has been largely refractive to the screens of temperature-sensitive mutants that have been employed for the isolation of the majority of the yeast splicing factors (for review see Hieter *et al.*, 1989).

The first gene, *SMD1*, encoding for a yeast Sm protein was discovered fortuitously near the *PRP38* gene (Rymond, 1993). The corresponding protein possesses 40% identity to the human D1 core protein. *In vivo* the lack of the yeast D1 is associated with a growth arrest and a splicing inhibition. Also, a decrease in the abundance of U1, U2, U4 and U5 snRNAs and a lack of cap modification is noticed, whereas U6 snRNA remains unaffected. The human D1 protein can rescue the lethal phenotype associated with the lack of the yeast D1, which confirms the homology between these proteins (Rymond *et al.*, 1993). In 1994, Fabrizio *et al.*, purified by immunoaffinity chromatography six potential yeast Sm proteins. These proteins possessed apparent molecular weights of 10, 11, 11.5, 12, 15 and 17 kD, however further identification has not been reported by these authors.

The yeast protein D3 was found by sequence similarity with the human SmD3 (Roy *et al.*, 1995). The yeast D3 possesses 51% identity to the human D3. It is essential for cell viability and splicing, and for the 5' cap modification of the U1, U2, U4, U5 snRNAs which become unstable in the cell when the production of D3 is shut down. The overproduction of D3 cannot compensate for the loss of D1, indicating that these proteins are functionally different (Roy *et al.*, 1995).

In an exhaustive two-hybrid screen using SmD3 as a bait, a protein named YER029cor SmB, which shares homology with the human B core protein, has been isolated and might interact with SmD3 (Fromont-Racine *et al.*, 1997).

Two additional potential yeast Sm proteins have been identified by computer sequence analysis (Séraphin, 1995). These proteins were named SmX2 and SmX3, and share

homology with the human SmG and F respectively (Séraphin, 1995). Both SmX2 and SmX3 are associated with the spliceosomal snRNAs.

The gene *SME1*, encoding for the yeast homologue of the human SmE protein, has been sequenced and characterised by Bordonné and Tarassov (1996). The corresponding protein is essential for cell viability and for U1, U2, U4 and U5 snRNAs stability. The human *SME* gene can complement a lethal phenotype associated with the disruption of the yeast *SME1,* which confirms the homology between the human and the yeast proteins.

The yeast homologue of the human SmD2 has been isolated during the purification of the U1 and the U2 snRNP particles (Neubauer *et al.*, 1997; Gottschalk *et al.*, 1998).

The investigation of Sm protein-protein interactions by directed two-hybrid tests and immunoprecipitation show that SmD3-B, E-G and F-E are in contact together (Camasses *et al.*, 1998).

*The Sm-like proteins of the U6 snRNP particle*. A list of U6 snRNP proteins so far identified is presented in Table 2B. The U6 snRNP particle does not share the common proteins with the other snRNPs. However this particle possesses a set of proteins which share homology with the Sm motif (reviewed in Fromont-Racine *et al.*, 1997; Mayes *et al.*, 1998). These proteins are classified as Sm-like polypeptides because they share homology with the Sm motifs, but unlike the core proteins, they are associated only with the U6 snRNA.

One of the U6 snRNP Sm-like protein, USS1 (also named SDB23), was initially isolated as a weak suppresser of a mutation affecting the function of a protein kinase involved in the yeast cell cycle (Parkes and Johnston, 1992). Further analyses revealed that USS1 was essential for cell viability and required for the U6 snRNA stability

(Cooper *et al.*, 1995 and 1995b). Sequence analysis indicated that USS1 shares significant homology with the human D proteins, in particular with D3.

A second gene named *SMX4* or *USS2*, has been identified by computer sequence analysis (Séraphin, 1995). The corresponding protein shares significant homology with the human SmD2 protein (Séraphin, 1995).

Computer analyses identified a set of putative Sm-like U6 snRNP proteins homologous to the yeast SmB, D1, E, F and G proteins (Fromont-Racine *et al.*, 1997). The ORFs encoding these proteins are named YJL124c, YBL026w, YER146w, YDR378c and YNL147w respectively. The corresponding proteins and also another one, Yjr022p, which contains an Sm site but does not share much homology with the cannonical Sm proteins, are associated with the U6 snRNA (Mayes *et al.*, 1998).

### *The snRNP specific proteins*

Unlike the core proteins, snRNP specific proteins are not required for the nuclear localisation of the snRNP particles. They are transported into the nucleus independently of any U snRNA through nuclear localisation signal sequences encoded within the proteins (Kambac *et al.*, 1992; Romac *et al.*, 1994).

Tables 2A and 2B show an exhaustive list of the human and yeast protein composition of the snRNP particles.

*The mammalian U1 snRNP specific proteins*. To date, three proteins have been isolated which belong to the human U1 snRNP particle. They are called U1-A, U1-70K and U1-C.

U1-A and U1-70K are members of a large family of RNA binding proteins that contain a conserved RNA recognition domain (RBD or RRM) (for description see section

**Table 2A: List of the U1 and U2 snRNP specific proteins in human and yeast.**

The size and the name(s) (when available) of these polypeptides are indicated. The corresponding human and yeast homologues are indicated by a dashed line.

## *SnRNP Specific Proteins*

| | Human | Yeast |
|---|---|---|
| **U1 snRNP** | 70 kD-70K | 34 kD-SNP1 |
| | 34 kD-A | 34 kD-MUD1 |
| | 22 kD-C | 27 kD-YU1C-YLR298 |
| | | 56 kD-SNU56-MUD10 |
| | | 57 kD-NAM8-MUD15 |
| | | 65 kD-SNU65-PRP42-MUD16 |
| | | 69 kD-PRP40 |
| | | 71 kD-SNU71 |
| | | 75 kD-PRP39 |

| | Human | | Yeast |
|---|---|---|---|
| **U2 snRNP** | **12S** | **17S** | |
| | 33 kD-A' | 33 kD-A' | |
| | 28.5 kD-B'' | 28.5 kD-B'' | 12.8 kD-YU2B" |
| | | 160 kD-SF3b160-SAP155 | |
| | | 150 kD-SF3b150-SAP145 | 50 kD-CUS1 |
| | | 120 kD-SF3b120-SAP130 | |
| | | 110 kD-SF3a120-SAP114 | 33 kD-PRP21-SPP91 |
| | | 92 kD | |
| | | 66 kD-SF3a66-SAP62 | 30 kD-PRP11 |
| | | 60 kD-SF3a60-SAP61 | 63 kD-PRP9 |
| | | 53 kD-SF3b53-SAP49 | 24 kD-HSH49 |
| | | 35 kD-SAP33 | |

**Table 2B: List of the U5, U6, U4/U6 snRNP and U4/U6.U5 tri-snRNP specific proteins in human and yeast.**

The size and the name(s) (when available) of these polypeptides are indicated. The names given by Mayes *et al.* (1998) are indicated by the symbol (ø). No human U6 snRNP proteins have been isolated to date. The yeast U6 snRNP proteins have been separated from the other snRNP proteins since to date, it is not known if some (or all) of them remain associated with the U6 snRNA in the U4/U6 snRNP and in the U4/U6.U5 tri-snRNP particles. Only Prp24p has been shown to promote U4/U6 snRNAs annealing and to interact transiently with the U4/U6 snRNP (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995; Ghetti *et al.*, 1995). U6 snRNP proteins sharing homology with the Sm motif of the core proteins are indicated with the symbol (℘).

Prp28p* is not stably associated with the tri-snRNP and is therefore not considered as a snRNP specific protein (Strauss and Guthrie, 1991).

Yjl124p* interacts loosely with the U6 snRNP (Mayes *et al.*, 1998).

## *SnRNP Specific Proteins*

**U4/U6.U5 tri-snRNP**

| | | Human | Yeast |
|---|---|---|---|
| **U5 snRNP** | **20S** | **25S** | |
| | 220 kD | 220 kD | 280 kD-PRP8 |
| | 200 kD | 200 kD | 246 kD-SNU246-BRR2-SLT22 |
| | 116 kD | 116 kD | 114 kD-SNU114 |
| | 110 kD | 110 kD | |
| | 102 kD | 102 kD | |
| | 100 kD | 100 kD | 67 kD-PRP28* |
| | 52 kD | 52 kD | |
| | 40 kD | 40 kD | |
| | 15 kD | 15 kD | |
| **U4/U6 snRNP** | 60 kD-SAP60 | 60 kD-SAP60 | 52 kD-PRP4 |
| | 90 kD-SAP90 | 90 kD-SAP90 | 56 kD-PRP3 |
| | | 61 kD | |
| | | 63 kD | |
| | | 27 kD | |
| | | 20 kD-SNUCYP-20 | |
| | | 15.5 kD | |
| | | | 28 kD-PRP38 |
| | | | 104 kD-PRP6 |

**Most similar to Sm proteins**

**Yeast U6 snRNP**

50 kD-PRP24
20 kD-SPB8$^{\theta}$-YJL124$^{\wp}$ * —— SmB
11 kD-USS6$^{\theta}$-YBL026-SMX5-SNP3 —— SmD1
10 kD-USS2-SMX4$^{\wp}$ —— SmD2
21 kD-USS1-YER112$^{\wp}$ —— SmD3
10.4 kD-USS7$^{\theta}$-YER146$^{\wp}$ —— SmE
14 kD-USS4$^{\theta}$-YDR378$^{\wp}$ —— SmF
12 kD-USS5$^{\theta}$-YNL147$^{\wp}$ —— SmG
14.5 kD-USS3$^{\theta}$-YJR022$^{\wp}$

1.4.7; for review see Birney *et al.*, 1993). They bind directly to the U1 snRNA stem-loop II and stem-loop I respectively (Scherly *et al.*, 1989; Query *et al.*, 1989). U1-70K has an extensive serine/arginine (SR) domain (see section 1.4.7 for description) and exists in multiple states of phosphorylation both *in vitro* and *in vivo* (Woppmann *et al.*, 1990, 1993; Wooley *et al.*, 1983). The phosphorylation state of U1-70K is essential for the splicing reaction but not for spliceosome formation (Tazi *et al.*, 1993b). Chemical protein-protein crosslinking studies have shown that U1-70K interacts via its N-terminus with the core proteins D2 and B' or B (Nelissen *et al.*, 1994).

In contrast to U1-A and U1-70K, the U1-C protein does not contain any recognised RNA binding motif, and depends on the other U1 snRNP proteins for its association with the U1 snRNP particle (Nelissen *et al.*, 1994). U1-C possesses a proline and methionine rich C-terminus, and a zinc-finger-like motif (see section 1.4.7 for description) at its N-terminus (Sillekens *et al.*, 1988). The zinc-finger-like motif is necessary and sufficient for the binding of U1-C to the U1 snRNP particle (Nelissen *et al.*, 1991). Chemical crosslinking experiments have demonstrated that U1-C interacts with the Sm B and B' proteins (Nelissen *et al.*, 1994). U1-C is needed for efficient interaction between the U1 snRNP and the pre-mRNA 5' splice site (Heinrich *et al.*, 1990; Nelissen *et al.*, 1991, 1994).

*The yeast U1 snRNP specific proteins*. The yeast homologues of the human U1-A and U1-70K, Mud1p and Snp1p respectively, have been isolated (Liao *et al.*, 1993; Smith and Barrell, 1991).

Mud1p (for mutant-u-die) was identified in a synthetic lethal screen with a mutant U1 snRNA. It is a non-essential protein, and its depletion is not associated with any detectable splicing defect, a further splicing decrease is observed only for introns which are normally inefficiently spliced. The absence of Mud1p slightly alters the mobility of the U1 snRNP. Mud1p contains two RNA recognition motifs, and is proposed to help the stabilisation of the U1 snRNA (Liao *et al.*, 1993). Both RRMs are important, with

the one located at the N-terminus being more likely to be involved in the association of Mud1p with the U1 snRNA, while the RRM located at the C-terminus may be involved in the splicing function of the protein (Tang and Rosbash, 1996). Chemical footprinting and nuclease S1 protection analyses indicate that Mud1p interacts with the yeast U1 snRNA stem-loop IIIc and its opposite single-stranded regions (region similar to the human stem-loop B) (Tang and Rosbash, 1996).

Snp1p shares 30% identity with the mammalian U1-70K protein (Smith and Barrell, 1991). It binds specifically and directly to the yeast U1 snRNA stem-loop I (Kao and Siliciano, 1992). A splicing defect is associated with the lack of Snp1p in cells, which can be complemented by a chimeric yeast-human Snp1p containing the central 188 amino acids of human U1-70K (Smith and Barrell, 1991). The carboxyl terminus of Snp1p does not share the SR domain found in metazoan U1-70K. This region is completely dispensable, while the amino terminus does not tolerate mutations or deletion (Hilleren *et al.*, 1995).

The gene encoding the yeast homologue of the human U1-C has recently been identified by data base searching, and was named YU1C (Tang *et al.*, 1997). Concomitantly, it has been isolated by purification and sequencing of the yeast U1 snRNP proteins (Neubauer *et al.*, 1997). The conservation between the yeast and human U1C protein is limited to the common N-terminal domain zinc finger-like region (Tang *et al.*, 1997). YU1C is essential for splicing, and its lack or the presence of a mutant protein affects the 5' end of the U1 snRNA which becomes more sensitive to RNase I digestion (Tang *et al.*, 1997).

The purification of the yeast U1 snRNP particle and its comparison with the human counterpart revealed that these two particles are different, the former being much larger (Fabrizio *et al.*, 1994). Further extensive analysis indicated that the yeast U1 snRNP contains nine specific proteins (Neubauer *et al.*, 1997; Gottschalk *et al.*, 1998), whereas the human counterpart possesses only three specific proteins (Lührmann *et al.*, 1990). Among these yeast U1 snRNP specific proteins, Prp39p (Lockhart *et al.*, 1994), Prp40p

(Kao and Siliciano, 1996) and Snu71p, Snu65p, Nam8p, Snu56p (Gottschalk *et al.*, 1998) have been isolated which lack any human homologues.

Prp39p is essential for splicing and is more likely to function at an early stage of spliceosome assembly. Prp39p is required to establish an efficient U1 snRNP/5' splice site interaction, and might also be involved in the release of the U1 snRNP particle from the spliceosome as spliceosome assembly proceeds (Lockhart and Rymond, 1994).

The gene encoding the Snu65 protein was isolated by others and named *PRP42* (McLean and Rymond, 1998). Prp42p is required for U1 snRNP biogenesis, since yeast strains depleted of Prp42p form incomplete U1 snRNPs that fail to produce stable complexes with pre-mRNA *in vitro* (McLean and Rymond, 1998).

Both Prp39p and Prp42p possess multiple copies of a variant tetratrico peptide repeat (TPR), an element implicated in a wide range of protein assembly events (Ordway *et al.*, 1995; Lamb *et al.*, 1995).

The gene encoding Prp40p was isolated in a genetic screen searching for suppressers of U1 snRNA cold-sensitive mutants (Kao and Siliciano, 1996). *PRP40* is essential for cell viability and the lack of the corresponding protein is associated with a splicing defect *in vitro*.

The genes encoding Snu56p, Nam8p and Snu65p were independently isolated in a synthetic lethal screen that causes lethality with otherwise viable mutations in U1 snRNA. They were named *MUD10*, *MUD15* and *MUD16* respectively (MUD for mutant-U1-die) (Gottschalk *et al.*, 1998). *SNU56/MUD10* is essential for cell viability. The corresponding protein exhibits some RNA-binding activity *in vitro,* despite the lack of any known RNA binding motifs (Gottschalk *et al.*, 1998). Snu56p is required for the stability and/or the formation of the commitment complex (Gottschalk *et al.*, 1998). Nam8p possesses three RNA binding motifs and is essential during meiosis and

for splicing of a meiosis-specific pre-mRNA (MER2 transcript) (Ogawa *et al.*, 1995; Nakagawa and Ogawa, 1997). This protein is probably involved in the stabilisation of U1 snRNA-5' splice site interactions (Gottschalk *et al.*, 1998). *SNU71* is essential for cell viability. This gene encodes a protein which contains several RS-repeats, but unlike the mammalian SR proteins these repeats are not concentrated within a domain but spread over the polypeptide chain (Gottschalk *et al.*, 1998). No known motifs of possible functional significance have been detected in SNU71.

*The mammalian U2 snRNP specific proteins*. Two distinct forms of the U2 snRNP particle have been isolated: a 12S form containing two U2 snRNP specific proteins denoted A' and B", and a 17S form which contains nine additional U2 snRNP specific proteins with molecular weights ranging from 35 kD to 160 kD. The 17S U2 snRNP is observed only under low salt conditions and appears to represent the functional form of this particle.

The U2-B" protein interacts directly with the U2 snRNA stem-loop IV, while A' appears to be associated by way of interaction with B" (Scherly *et al.*, 1990).

Most of the 17S U2 snRNP specific proteins interact with the 5' half of the U2 snRNA (Behrens *et al.*, 1993). These proteins have been the most studied of the snRNP specific proteins (reviewed in Hodges and Beggs, 1994; Krämer, 1995), and several names have been given for identical proteins isolated by different groups. At least seven of the 17S particle specific proteins have been isolated in HeLa cells as two tightly associated protein complexes named SF3a and SF3b. These two subunits are required for the assembly of the mature U2 snRNP particle and for the targeting of this snRNP to the pre-mRNA. SF3b binds to the 12S U2 snRNP particle (Lührmann *et al.*, 1990) and the resulting 15S particle is converted into the active 17S U2 snRNP after interaction with SF3a. SF3a does not associate by itself with the 12S particle. It is composed of three associated subunits of 60 kD (SF3a60p), 66 kD (SF3a66p) and 120 kD (SF3a120p) (Brosi *et al.*, 1993a, 1993b). These same proteins were also named Sap61, 62, 114

(spliceosome associated proteins) respectively (Bennett *et al.*, 1992, Chiara *et al.*, 1993). Protein-protein interactions detected between the human U2 snRNP specific proteins are similar to those observed for their yeast homologues. Sap62p (human homologue of the yeast Prp11p) interacts with Sap114p (human homologue of the yeast Prp21p) (Bennett and Reed, 1993) and Sap61p (human homologue of the yeast Prp9p) interacts with Sap114p (Chiara *et al.*, 1994). SF3b is composed of five proteins named Sap33p, 49, 130, 145, 155 (for review see Hodges and Beggs, 1994). With the exception of Sap33p, these proteins were found to be in close proximity to the 3' extremity of the nuclear pre-mRNA intron in the prespliceosome complex (Staknis and Reed, 1994). Behrens *et al.* (1993) have isolated nine proteins associated with the 17S complex, possessing apparent molecular weights of 35, 53, 60, 66, 92, 110, 120, 150 and 160 kD. The 35, 53, 120, 150 and 160 kD polypeptides have been reported as SAP33, 49, 130, 145 and 155 respectively by Bennett *et al.*, 1992, 1993 and Staknis and Reed, 1994. They have been isolated also by Brosi *et al.* (1993) and named SF3b53, 120, 150, and 160 according to their molecular weights, while the 35 kD polypeptide did not appear to be part of the SF3b complex. The 92 kD protein has been detected only by Behrens *et al.* (1993), and its presence in the U2 snRNP particle needs confirmation. The protein-protein interactions occuring in this protein network are not very well known. It has been demonstrated that Sap49p and Sap145p interact together and with the pre-mRNA and may be involved in tethering the U2 snRNP to the branchsite (Champion-Arnaud and Reed, 1994).

*The Yeast U2 snRNP specific proteins*. *MSL1* (*MUD* synthetic lethal 1), encodes the yeast homolog (YU2B") of the mammalian U2B" (Tang *et al.*, 1996). The YU2B" protein is much smaller that U2B" and it contains only one RBD at its N-terminus (Tang *et al.*, 1996).

Several lines of evidence have shown that the yeast proteins Prp9p and Prp11p are the homologues of the human SF3a60p (Sap61p) and SF3a66p (Sap62p) respectively. Both function during prespliceosome assembly in yeast, and interact genetically and

physically with a third protein named Prp21p (or Spp91p) (Behrens *et al.*, 1993; Legrain *et al.*, 1993, 1993b; Chapon and Legrain, 1992). Prp21p contains significant homology to the human SF3a120 protein. Prp9p possesses 30% identity to the human SF3a60p, the highest homology being present in a zinc finger-like region located at the C-terminus of both proteins. The Prp9p zinc finger-like motif has been replaced by the equivalent domain of mammalian SF3a60p, the resulting chimeric polypeptide was able to rescue the temperature-sensitive phenotype of a *PRP9-1* mutant strain, suggesting that the structure and the function of this domain were evolutionarily conserved (Krämer *et al.*, 1994). Using the two-hybrid system, Legrain and Chapon (1993b) show that the carboxyl terminus of Prp11p binds to Prp21p. Also, Prp21p binds to Prp9p and to Prp11p forming a bridge between them (Legrain and Chapon, 1993b) while no interaction was detected between Prp9p and Prp11p. Prp9p, Prp11p and Prp21p interact with a region of the U2 snRNA including the stem-loop IIa. Like their human homologues, they mediate the association of the U2 snRNP with the pre-mRNA (Wells and Ares, 1994).

To date, only two yeast homologues of the human SF3b proteins have been isolated (Cus1p and Hsh49p) (Wells *et al.*, 1996). *CUS1* (cold-sensitive U2 suppressor) has been isolated as a suppresser of a cold-sensitive mutation associated with a mutation in U2 snRNA stem-loop IIa (Wells *et al.*, 1996). Cus1p is essential and shares significant homology with Sap145p, however, it is much smaller (Wells *et al.*, 1996). Hsp49p (human SAP homologue 49) has been identified by computer sequence analysis as a putative homologue of the human Sap49p (Igel *et al.*, 1998) Further investigations are required to determine if Hsh49p is the yeast counterpart of Sap49p. Interestingly, in an exhaustive two-hybrid screen using either Cus1p or Yor319p (other name for Hsh49p) as baits, Fromont-Racine *et al.* (1997) found that the two polypeptides interact together, and with other proteins.

*The mammalian U5 snRNP specific proteins.* The 20S U5 snRNP particle contains at least nine specific proteins (Bach *et al.*, 1989). These proteins are retained following the formation of the U4/U6.U5 tri-snRNP complex. At least one of them, p220, remains in the spliceosome and is suspected to play a role in modulating or stabilising RNA-RNA contacts between the U5 snRNA and the pre-mRNA (for review see Newman, 1997). p220 is the human homologue of the yeast Prp8p (Garcia-Blanco *et al.*, 1990). UV-crosslinking experiments have demonstrated that p220 interacts with the pre-mRNA (Garcia-Blanco *et al.*, 1990) and it is in close proximity to the 5' splice spite before the first step of splicing reaction (Wyatt *et al.*, 1992).

The 20S U5 snRNP particle possesses an RNA-dependent ATPase activity associated, at least, with two polypeptides named U5-200 kD and U5-100 kD (Laggerbauer *et al.*, 1996).

U5-200 kD possesses two conserved domains characteristic of the DEXH-box protein family of RNA-stimulated ATPases and putative RNA helicases (see section 1.4.7 for description) (Lauber *et al.*, 1996). The depletion of U5-200 kD in a HeLa cell extract blocks the second step of the splicing reaction, suggesting that this protein probably remains in the spliceosome and may be involved in the second step of splicing (Lauber *et al.*, 1996). This polypeptide is the first putative RNA helicase shown to be involved in splicing in humans.

The U5-100 kD polypeptide has been recently characterised (Teigelkamp *et al.*, 1997). This protein is a member of the DEAD-box family of ATP-dependent putative RNA helicases. In addition, it harbours at its N-terminal an RS domain. Protein analysis by two-dimensional gels revealed that U5-100 kD exists in three different stages of phosphorylation. Sequence analysis indicated that U5-100 kD is the human homologue of the yeast Prp28p, but its role in splicing remains unknown (Teigelkamp *et al.*, 1997). The yeast Prp28p is also a member of the DEAD-box family and is required for the first step of splicing (Strauss *et al.* 1991). Unlike its human homologue, Prp28p is not stably

associated with snRNP particles and therefore is not considered a snRNP specific protein (Strauss and Guthrie, 1994).

The U5-116 kD polypeptide has been recently characterised (Fabrizio *et al.*, 1997). It contains a consensus sequence (G domain) involved in the binding and hydrolysis of GTP. UV-crosslinking experiments confirmed that this polypeptide specifically binds to GTP, and that this GTP-association was required for the protein function (Fabrizio *et al.*, 1997).

*The yeast U5 snRNP specific proteins*. Prp8p (Lossky *et al.*, 1987) is the yeast homologue of p220 (Garcia-Blanco *et al.*, 1990). This protein is exceptionally large (280 kD) and its size and sequence are highly conserved (Hodges *et al*, 1995). Prp8p is associated with the U5 snRNP and the U4/U6.U5 tri-snRNP (Lossky *et al.*, 1987; Brown and Beggs, 1992). It is required for the stable formation of the tri-snRNP and for its association with the spliceosome (Brown and Beggs, 1992). Prp8p remains in the spliceosome (Whittaker *et al.*, 1991; for review see Beggs *et al.*, 1995) and stabilises the interactions between U5 snRNA and the 5' and 3' splice sites, to anchor and align them in the reactive centre of the spliceosome (Teigelkamp *et al.*, 1995; Dix *et al.*, 1998). It is the only yeast splicing factor that has been demonstrated to affect both chemical steps of splicing (Jackson *et al.*, 1988; Umen and Guthrie, 1995).

The gene encoding the yeast homologue of the human U5-200 kD protein was found by computer data base search, and the corresponding polypeptide was named Snu246p (Lauber *et al.*, 1996). This yeast protein is essential for cell viability and required for pre-mRNA splicing *in vivo*. Immunoprecipitation of an HA-epitope tagged Snu246p showed that it was associated with the tri-snRNP particle (Lauber *et al.*, 1996). The same gene has been identified concomitantly by different groups and named differently: *BRR2* (for bad response to refrigeration) (Noble and Guthrie,1996) was isolated as a cold-sensitive mutant; *SLT22* (for synthetic lethal with U2) was isolated in a genetic screen as a synthetic lethal with a mutant U2 snRNA (*snr20-11nt*) which potentially

disrupts U2/U6 snRNAs Helix II interaction (Xu *et al.*, 1996); *RSS1* was isolated in a genetic screen suppressing a splicing defect of an actin pre-mRNA-*lacZ* fusion construct, this defect was caused by the addition of an artificial secondary RNA structure in exon 2 adjacent to the 3' splice site (Lin and Rossi, 1996). This protein belongs to the DEXH ATPase/RNA helicases family (distantly related to the ATPase/RNA helicases of the DEAD and DEAH family) (Noble and Guthrie, 1996; Xu *et al.*, 1996). Slt22p is essential for cell viability and exhibits an ATPase activity which is preferentially stimulated by annealed U2/U6 snRNAs (Xu *et al.*, 1996). Immunoprecipitation analyses have shown that Brr2p is associated with the pre-mRNA and splicing intermediates (Raghunathan and Guthrie, 1997). The same authors demonstrated that Brr2p is indeed an RNA helicase which dissociates U4 snRNA from the U6 snRNA in an ATP-dependent fashion.

The yeast homologue of the human U5-116 kD protein was identified and named Snu114p (Fabrizio *et al.*, 1997). Snu114p is essential for cell viability and splicing *in vivo*.

*The yeast U6 snRNP particle specific proteins*. As yet, no human U6 snRNP specific protein has been identified. In yeast, in addition to the Sm-like proteins (see section 1.4.3.2 for description), one U6 snRNP specific protein (Prp24p) has been identified (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995; Ghetti *et al.*, 1995; Raghunathan and Guthrie, 1998). Prp24p was initially isolated as a suppresser of a single point mutation (G14C) in the U4 snRNA. This particular mutation destabilises the U4/U6 snRNAs helix II and is associated with a cold-sensitive phenotype (Shannon and Guthrie, 1991). Prp24p contains three RNA-binding motifs and it directly contacts the U6 snRNA probably promoting the formation of the U4/U6 snRNP particle (Shannon and Guthrie, 1991). Prp24p may also be involved in the dissociation of U4/U6 snRNP prior to the first step of splicing. It has been proposed to distort helix II in the U4/U6 snRNAs partial hybrid helping another splicing factor (maybe an RNA helicase) to unwind this structure (Jandrositz and Guthrie, 1995; Ghetti *et al.*, 1995).

Prp24p is involved in the recycling of U4 and U6 snRNPs and reanneals U4/U6 snRNAs after the splicing reaction (Raghunathan and Guthrie, 1998). To date, Prp24p is the only splicing factor shown to contact directly the U6 snRNA (Jandrositz and Guthrie, 1995; Ghetti *et al.*, 1995). A genetic interaction between Prp24p and Prp21p has been reported, suggesting a link between the U2 snRNP and the U6 snRNP (Vaidya *et al.*, 1996).

*The mammalian U4/U6 snRNP specific proteins.* In the human 10S U4/U6 snRNP particle two proteins (60 kD and 90 kD) are present in addition to the Sm proteins. They are named Sap60p and Sap90p respectively (Gozani *et al.*, 1994), and are the human homologues of the yeast Prp4p and Prp3p respectively (Lauber *et al.*, 1997). Antibodies raised against the 60 kD and the 90 kD proteins specifically recognise the U4/U6 snRNP and the tri-snRNP particles (Lauber *et al.*, 1997).

Like its yeast counterpart, the human 60 kD protein possesses at its C-terminus WD repeats, which are essential for its function during splicing (Hu *et al.*, 1994; Ayadi *et al.*, 1997).

Not much is known about the 90 kD polypeptide. It possesses some similarity to a double-stranded RNA binding motif.

*The yeast U4/U6 snRNP specific proteins*. Two U4/U6 snRNP specific proteins, namely Prp4p and Prp3p, have been identified genetically (Banroques *et al.*, 1989; Petersen-Bjørn *et al.*, 1989).

Prp4p is required for the association of the U4/U6 snRNP with the U5 snRNP, and interacts with the conserved 5' stem-loop of the U4 snRNA (Banroques *et al.*, 1989; Petersen-Bjørn *et al.*, 1989; Bordonné *et al.*, 1990; Xu *et al.*, 1990).

Prp3p was proposed to be U4/U6 snRNP-specific based on the observation that overproduction of Prp3p can compensate for a defective Prp4 protein (Last *et al.*, 1987). Further characterisation demonstrated that Prp3p was required for spliceosome formation (Anthony *et al.*, 1997). It is needed for efficient assembly and/or stability of the U6 snRNP and U4/U6 snRNP particles (Anthony *et al.*, 1997). Antibodies raised against Prp3p recognise specifically the U4, U5 and U6 snRNAs, and only U4 and U6 snRNA at higher salt concentrations, suggesting that this protein is associated with the U4/U6 snRNP and the tri-snRNP particles.

*The mammalian U4/U6.U5 tri-snRNP specific proteins*. In addition to the aforementioned 20S U5 snRNP and 12S U4/U6-specific proteins, the 25S tri-snRNP contains five other specific proteins which possess apparent molecular weights of 15.5, 20, 27, 61 and 63 kD (Behrens *et al.*, 1991).

The 20 kD protein (also named SNU-CYP-20) has been recently purified and characterised (Teigelkamp *et al.*, 1998). It shares significant homology with the cyclophilin protein family. Cyclophilins possess a peptidyl-prolyl *cis-trans* isomerase activity (PPIase) (see section 1.4.7 for description) which may be involved in protein conformation changes. Further investigations indicated that this polypeptide interacts with at least the 60 kD protein (Teigelkamp *et al.*, 1998).

The 27 kD contains at the N-terminus a region enriched in arginine and serine residues (SR domain). In contrast with the known SR proteins, it lacks any similarity with any of the characterised RNA binding domains (Fetzer *et al.*, 1997). This protein can be differentially phosphorylated *in vitro* by a snRNP associated kinase activity (as yet uncharacterised) (Fetzer *et al.*, 1997).

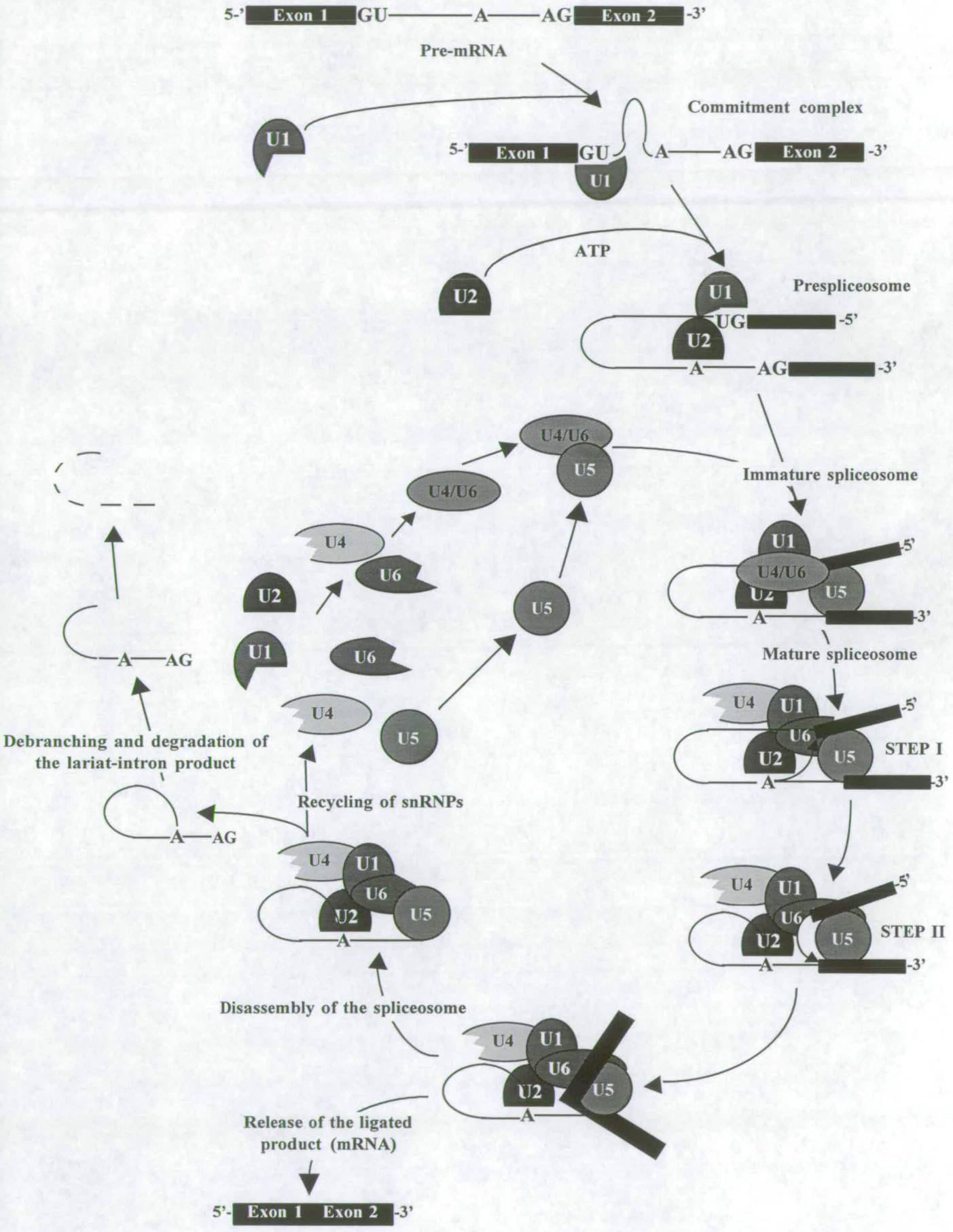
*The yeast tri-snRNP specific proteins*. The yeast and the human tri-snRNPs possess structural similarities which probably reflects similarities between the corresponding splicing machinery (Fabrizio *et al.*, 1994).

The yeast Prp6p is associated with the U4, U5 and U6 snRNAs and is required for the stability of the U4/U6.U5 tri-snRNP particle (Galisson and legrain, 1993).

Prp38p is a component of the tri-snRNP (Xie *et al.*, 1998). Its exact function remains speculative. It might be involved in the dissociation of U4/U6 snRNAs interactions as the tri-snRNP enters the prespliceosome (Xie *et al.*, 1998).

### 1.4.4 U2-Spliceosome assembly pathway

The spliceosome assembly pathway is shown in Figure 7. It starts by the binding of the U1 snRNP to the pre-mRNA in an ATP-independent fashion (Rosbash and Séraphin, 1991). This binding commits the pre-mRNA to the splicing pathway and corresponds to the formation of the commitment complex (complex E in mammals). Subsequently the U2 snRNP associates with the branchsite to form the prespliceosome complex (complex A in mammals) (Konarska and Sharp, 1986; Zillmann *et al.*, 1988). This interaction, as well as subsequent steps of spliceosome assembly, is ATP-dependent. In the next step U4/U6 snRNP and U5 snRNP which are assembled as a U4/U6.U5 tri-snRNP associate with the prespliceosome, resulting in the formation of an immature complex (complex B in mammals) (Lossky *et al.*, 1987; Pikielny *et al.*, 1896; Cheng and Abelson, 1987). After conformational rearrangements, the immature spliceosome is converted into an active complex (complex C in mammals). Thus splicing occurs, followed by the release of the newly joined exons and dissociation of the spliceosome, then degradation of the intron-lariat product (Konarska and Sharp, 1987; Cheng and Abelson, 1987; Blencowe *et al.*, 1989). The very long half-life of the snRNAs suggests that they are recycled and enter the spliceosome assembly pathway repeatedly *in vivo*.


5-' Exon 1 GU A AG Exon 2 -3'
Pre-mRNA
U1
Commitment complex
5-' Exon 1 GU A AG Exon 2 -3'
U1
ATP
U2
Prespliceosome
U1
UG -5'
U2
A AG -3'
U4/U6
U5
Immature spliceosome
U4/U6
U4
U6
U2
U1
U6
U5
U1
U4/U6
U5
U2
A -3'
5'
Mature spliceosome
A AG
U4
U5
U4 U1
U6
U2 U5
STEP I
A -3'
Debranching and degradation of the lariat-intron product
Recycling of snRNPs
A AG
U4 U1
U6
U2 U5
A
U4 U1
U6
U2 U5
STEP II
A -3'
Disassembly of the spliceosome
U4 U1
U6
U2 U5
A
Release of the ligated product (mRNA)
5'- Exon 1 Exon 2 -3'

#### 1.4.4.1 The commitment complex

It was originally defined in yeast nuclear extracts (Legrain *et al.*, 1988; Séraphin and Rosbash, 1989). Two complexes (CC1 and CC2) can be distinguished on a non-denaturing gel electrophoresis system (Séraphin and Rosbash, 1989). The formation of CC1 requires only the 5' splice site sequence, whereas the formation of CC2 requires both the 5' splice site and the branchsite sequences (Séraphin and Rosbash, 1991). CC1 and CC2 assemble in an ATP-independent fashion and contain the U1 snRNP.

A similar ATP-independent complex has been purified in mammalian extracts and was named complex E (Michaud and Reed, 1991). The formation of this complex requires, in addition to the 5' splice site and the branchpoint, the polypyrimidine tract (Jamison *et al.*, 1992). Unlike CC1 and CC2, complex E contains both U1 and U2 snRNPs, although the U2 snRNP is weakly associated (Hong *et al.*, 1997).

Both the human and the yeast commitment complexes involve base-pairing of the 5' end of the U1 snRNA to the 5' splice site region of the pre-mRNA (Zhuang and Weiner, 1989; Séraphin *et al.*, 1988; Siliciano and Guthrie, 1988). This interaction is facilitated by several polypeptides.

***In mammals***

The U1 snRNP specific 70 kD and C proteins and the SR protein ASF (alternative splicing factor; also named SF2 for splicing factor 2) promote these contacts. ASF/SF2 recruits the U1 snRNP to the 5' splice site by interacting simultaneously with the pre-mRNA and the U1-70K protein (kohtz *et al.*, 1994), while the splicing factor U2AF (U2 snRNP auxiliary factor) initiates the recognition of the 3' splice site (Zuo and Maniatis, 1996; Eperon *et al.*, 1993; Kohtz *et al.*, 1994; Staknis and Reed, 1994).

U2AF is a heterodimer composed of two subunits of 65 kD and 35 kD (named U2AF65 and U2AF35 respectively), which are both required for splicing (Zamore and Green, 1989). U2AF65 binds to the 3' splice site region prior to U2 snRNP addition (Ruskin *et al.*, 1988; Zamore and Green, 1989). U2AF65 interacts with the SF1 splicing factor (Arning *et al.*, 1996) and positions it to the branchpoint (Abovich and Rosbash, 1997; Berglund *et al.*, 1997). U2AF35 possesses at its amino-terminal an RS domain, used to mediate interaction with other SR proteins (Zuo and Maniatis, 1996), while its carboxyl-terminal domain is required for formation of U2AF35/65 heterodimer (Zamore *et al.*, 1992; Zhang *et al.*, 1992). Thus U2AF35 functions as a bridge between U2AF65 and proteins containing one or more RS domain(s) (Reed, 1996). U2AF35 has also been proposed to link the 3' and the 5' splice sites by interacting with SR proteins, while SR proteins interact with the U1 snRNP U1-70K protein (Wu and Maniatis, 1993). During the transition from commitment complex to prespliceosome, U2AF65 is phosphorylated and becomes less tightly bound to the pre-mRNA. U2AF65 remains loosely associated to the pyrimidine tract at least through complex-B assembly (Champion-Arnaud and Reed, 1994).

***In yeast***

Prp39p appears to facilitate the association of the U1 snRNP with the 5' splice site (Lochbart and Rymond, 1994; Kao and Siliciano, 1996).

The yeast gene encoding for the homologue of U2AF65 was isolated in a synthetic lethal screen with U1 snRNA and was named *MUD2* (Abovich *et al.*, 1994). Like its metazoan counterpart, Mud2p interacts with the pre-mRNA prior to the addition of the U2 snRNP and its association with the commitment complex requires the branchpoint sequence. Mud2p mediates the association of the U1 snRNP and the 3' splice site region.

Genetic experiments indicate that Mud2p is associated with the U1 snRNP protein Mud1p and contacts the U2 snRNP protein Prp11p, suggesting that Mud2p is at the U1-U2 snRNP interface during spliceosome formation (Abovich *et al.*, 1994).

In yeast, few proteins containing a classical RS domain have been isolated. The SR dipeptides are scattered through the polypeptide (Abovich *et al.*, 1994), therefore it is difficult to propose a picture of interactions between SR proteins that bridge the 5' and the 3' splice sites. However recently, the gene encoding a protein named BBP (branchpoint bridging protein) also named Msl5p (MUD synthetic-lethal 5p), has been isolated in a synthetic-lethal screen with *MUD2* (Abovich and Rosbach, 1997). BBP is the yeast homologue of the mammalian SF1 splicing factor (Abovich and Rosbash, 1997). BBP interacts with Mud2p and with the U1 snRNP specific protein Prp40p thereby bridging the 5' and the 3' splice sites (Abovich and Rosbash, 1997). BBP directly contacts the pre-mRNA near the branchpoint sequence during commitment complex assembly (Berglund *et al.*, 1997). Further investigations indicated that BBP is the yeast homologue of the human SF1 which interacts with U2AF65 (Arning *et al.*, 1996; Abovich and Rosbash, 1997).

It has been shown that in mammals the $m^7G$ cap binding proteins CBP20 and CBP80 enhance the U1 snRNP-5' splice site interaction, at least in the case of cap-proximal 5' splice sites (Lewis *et al.*, 1996a). A similar role in early splicing complex formation was also demonstrated for the yeast homologues of CBP20 and CPB80 (Mud13p and yCbp80p, respectively) (Colot *et al.*, 1996; Lewis *et al.*, 1996b). This is consistent with reports suggesting that a $m^7G$ cap enhances yeast pre-mRNA splicing *in vivo* (Schwer and Shuman, 1996; Fresco and Buratowski, 1996).

### 1.4.4.2 The prespliceosome or complex A

The recognition of the branchpoint sequence is essential for the assembly of the complex A. A short duplex is formed between the branchpoint sequence and U2 snRNA (Madhani and Guthrie, 1994c) which is thought to bulge the branchsite adenosine residue that initiates the first trans-esterification reaction (for review see Moore *et al.*, 1993).

***In mammals***

Stabilisation of this U2 snRNA-branchpoint interaction is mediated by the binding of several U2 snRNP proteins to a 20-nt region upstream of the branchpoint sequence which is designated the anchoring site (Gozani *et al.*, 1996).

Among these polypeptides, the mammalian U2 snRNP specific subunits SF3a and SF3b have been UV-crosslinked to the pre-mRNA (with the exception of the SAP 130 polypeptide) (Brosi *et al.*, 1993a, 1993b; Champion-Arnaud and Reed, 1994; Gozani *et al.*, 1996), U2AF65 interacts directly with the polypyrimidine tract and lies adjacent to the branchpoint sequence (Zamore and Green, 1989; Singh *et al.*, 1995).

Until recently, the ATP requirement for this step was not clear, since none of the known mammalian factors necessary for complex A assembly has the sequence characteristic of proteins that hydrolyse ATP. However recently, a protein, named UAP56 (for U2 associated protein), was identified (Fleckner *et al.*, 1997). This polypeptide is a member of the DEAD box family of RNA-dependent ATPases. It was proposed to stabilise the U2 snRNP-branchpoint interaction and to be responsible for the ATP-dependency of complex A formation.

### *In yeast*

The composition of the prespliceosome is less well defined. It contains at least Prp5p, Prp9p, Prp11p and Prp21p. These polypeptides are required for U2 snRNP binding to the U1 snRNP-containing complex (Ruby *et al.*, 1993; Abovich *et al.*, 1990; Legrain and Chapon, 1993b; Legrain *et al.*, 1993; Chapon and Legrain, 1992; Arenas and Abelson, 1993). Prp21p binds directly to Prp9p and Prp11p acting as a bridge. *PRP5* interacts genetically with *PRP9*, *PRP21* and *PRP11* as well as with U2 snRNA.

Prp5p is a member of the DEAD-box family of proteins (Dalbadie-McFarland and Abelson, 1990), and possesses an RNA-dependent ATPase activity (O'Day *et al.*, 1996). In particular, the ATPase activity of Prp5p is stimulated by the yeast U5 and U2 snRNAs with U2 snRNA being the most effective, suggesting that U2 snRNA may be a natural substrate for this putative RNA helicase. Its association with the commitment complex is ATP-independent, but its activity is likely to require ATP-hydrolysis and to be responsible (maybe just partially) for the ATP-dependence of the prespliceosome assembly step. Prp5p is required for U2 snRNP entry (Ruby *et al.*, 1993; Wells and Ares, 1994).

#### 1.4.4.3 Addition of the U4/U6.U5 tri-snRNP

The association of the tri-snRNP with the prespliceosome has been less well characterised.

### *In mammals*

Gradient-purified pre-spliceosomes can be chased into spliceosomes by addition of phosphorylated recombinant SR proteins (Roscigno and Garcia-Blanco, 1995). Thus SR proteins are proposed to escort the tri-snRNP as it associates with the prespliceosome complex. Interestingly the analysis of the tri-snRNP specific 27 kD protein and the U5-snRNP specific 100 kD protein (which remains in the tri-snRNP) indicated that both contain an RS domain and were capable of being phosphorylated by

a snRNP-associated kinase activity (Fetzer *et al.*, 1997; Teigelkamp *et al.*, 1997). These proteins may interact with other SR proteins, leading to the association of the tri-snRNP particle with the prespliceosome.

***In yeast***

*PRP31* encodes an essential factor required for the transition of prespliceosome to spliceosome complex (Weidenhammer *et al.*, 1997), probably by mediating or stabilising the association of the tri-snRNP with the prespliceosome (Weidenhammer *et al.*, 1997).

Prp6p is also required for tri-snRNP assembly and addition to the prespliceosome (Galisson and Legrain, 1993).

Prp38p is required for the maturation of the spliceosome. It may facilitate the unwinding of the U4 /U6 snRNA interactions within the tri-snRNP particle, to free the U6 snRNA and allow maturation of the spliceosome (Xie *et al.*, 1998).

#### 1.4.4.4 Structural rearrangements of snRNAs toward formation of the active spliceosome

Numerous conformational changes occur during spliceosome assembly. On the basis of compensatory mutation studies carried out in yeast, another conformation of the U6 snRNA has been proposed in which U6 interacts with U2 snRNA in the spliceosome (Madhani and Guthrie, 1992). This U6-U2 structure is mutually exclusive with the U4-U6 base-pairing and is therefore assumed to occur in the active spliceosome after, or concomitant with U4 and U6 dissociation (Wassarman and Steitz, 1992).

As tri-snRNP enters the spliceosome, destabilisation of the extensive base-pairing interaction between the U4 and U6 snRNAs occurs (Pikielny *et al.*, 1986; Cheng and

Abelson, 1987; Lamond *et al.*, 1988). U4 snRNA appears to remain within the splicing complex, but its presence is not essential for the subsequent catalytic steps (Yean and Lin, 1991). However, depending on the isolation conditions of the spliceosomal complexes, the U4 snRNA may be lost (Pikielny *et al.*, 1986; Cheng and Abelson, 1987; Konarska and Sharp, 1987). As U6 snRNA dissociates from the U4 snRNA, it starts to interact with the 5' splice site (Li and Brow, 1996) replacing the U1 snRNA (Konforti and Konarska, 1994; Lesser and Guthrie, 1993; Sontheimer and Steitz, 1993; Kandel-Lewis and Séraphin, 1993; Sawa and Abelson, 1992; Sawa and Shimura, 1992).

U1 snRNA is not required for the catalysis of the splicing reaction (Yean and Lin, 1996). At least in HeLa cells, the U1 snRNP remains associated with spliceosome complexes E, A, B and with the mature spliceosome (Ast and Weiner, 1996, 1997). However it is predominantly detected in the commitment complex (Bennett *et al.*, 1992) and its presence is weaker as U2 snRNP and U4/U6.U5 tri-snRNP assemble (Bennett *et al.*, 1992).

In the active spliceosome the U1 snRNA is supplanted by the U5 and the U6 snRNAs at the pre-mRNA 5' splice site (for review see Newman, 1994; Ares and Weiser, 1995). It has been suggested that the base-pairing between U1 snRNA and the pre-mRNA is disrupted even earlier, before U2 snRNA binds (Liao *et al.*, 1992; Champion-Arnaud *et al.*, 1995). U6 snRNA thus forms a new duplex with the U2 snRNA (Madhani and Guthrie, 1992), while the U5 snRNA interacts with a small number of poorly conserved exon nucleotides at both the 5' and the 3' splice sites (Madhani and Guthrie, 1994c). These interactions are stabilised by the Prp8 protein (or the 220 kD protein in human) (for review see Beggs *et al.*, 1995; Dix *et al.*, 1998).

### 1.4.4.5 The catalytic centre of the spliceosome

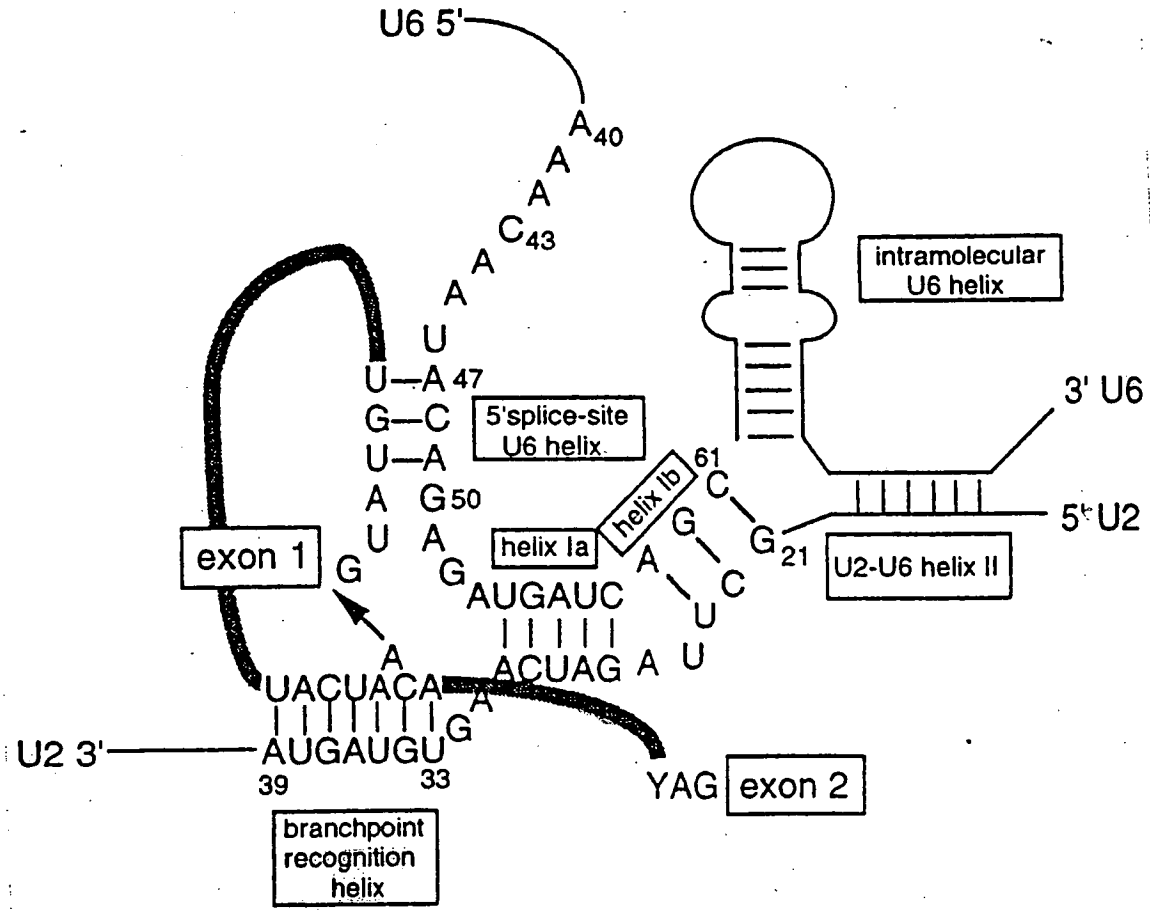
A representation of some RNA-RNA interactions involved in the organisation of the catalytic centre of the spliceosome is shown in Figure 8. Base-pairing interaction between U2 and U6 snRNAs, in conjunction with the U6 snRNA-5' splice site duplex, are though to position the branchpoint sequence near the 5' splice site to catalyse the first step of the splicing reaction (Sun and Manley, 1995; for review see Madhani and Guthrie, 1994c). Mutations in either the U6 snRNA or the U2 snRNA portion of U6-U2 helix I result in first or second step splicing defects (Fabrizio and Abelson, 1990; Madhani and Guthrie 1992; McPheeters and Abelson, 1992). A tertiary interaction between U6-G52 and the helix I (U2-A25) bulge has been demonstrated (Madhani and Guthrie, 1994a). This tertiary interaction might function in a network of RNA-RNA contacts that together form an active site involved in 3' splice site cleavage and exon ligation (Madhani and Guthrie, 1994a).

The highly conserved loop I of U5 snRNA interacts with the 5' exon before the first step of pre-mRNA splicing and with both the 5' and 3' exons following the first step (Newman and Norman, 1991, 1992). U5 snRNA does not make an essential contribution to the selection of the 5' splice site but its loop I is important for the second step of the splicing reaction (O'Keefe *et al.*, 1996; O'keefe and Newman, 1998). The interactions between the 5' and 3' exons and U5 snRNA loop I probably hold the exons in the correct orientation for the second step of splicing. Prp8p/p220 has been implicated in the recognition of both the 5' and the 3' splice sites (Reyes *et al.*, 1996; Umen and Guthrie, 1995, 1996) and is closely associated with the intron branchsite in active spliceosomes (MacMillan *et al.*, 1994). This raises the intriguing possibility that Prp8p/p220 might contribute functional groups to the catalytic centre of the spliceosome, in addition to assisting in the precise positioning of the U5 snRNA terminal loop at the 5' splice site (reviewed in Newman, 1997).

Other essential features contributing to splicing are the phosphate-ribose backbone and the 2'-hydroxyl groups of the U6 snRNA molecule. The defects caused by site-specific

**Figure 8. Schematic representation of the U2/U6 snRNA interactions in the catalytic centre of the yeast spliceosome (Madhani and Guthrie, 1992).**

Base-pairing between the stem I of U6 snRNA and U2 snRNA (helix I) allows the stem II region of U6 snRNA to refold into the same intramolecular helix present in the singular U6 form. In yeast, Helix Ib and II are functionally redundant, each being necessary only when the other is disrupted (Field and Friesen, 1996).


U6 5'
A40
A
A
C43
A
A
U
U—A 47
G—C
U—A
A G50
U A
G
5'splice-site
U6 helix
intramolecular
U6 helix
3' U6
5' U2
61
C
G
G21
helix Ib
helix Ia
U2-U6 helix II
exon 1
G
A
AUGAUC
ACUAG A U
A
C
U
UACUACA
A
G
U2 3'
AUGAUGU
39
33
YAG
exon 2
branchpoint
recognition
helix

substitution of phosphate by thiophosphate groups, or incorporation of deoxyribonucleotides in the U6 snRNA have implicated particular phosphates or hydroxyl groups of the U6 snRNA in the catalysis of splicing (Fabrizio and Abelson, 1992; Kim *et al.*, 1997).

The spliceosomal snRNAs are though to function as catalysts for the two reactive steps of pre-mRNA splicing. This hypothesis is founded on the knowledge that group I and group II introns are capable of selfsplicing (see sections for description), and the observation of several similarities between some snRNA-snRNA and snRNA-pre-mRNA interactions and some features of group I and group II introns involved in splicing. However snRNAs are not able to drive splicing of nuclear introns by themselves; proteins are required for spliceosome assembly and splicing reaction.

#### 1.4.4.6 Spliceosome disassembly

This is an active process that requires ATP and protein factors (Company *et al.*, 1991; Sawa and Shimura, 1991). In particular, the release of the spliced exons requires an RNA-dependent ATPase named Prp22p (Company *et al.*, 1991). Prp22p is a putative RNA helicase which is required for spliceosome dissociation. *In vitro* assays show that Prp22p is an RNA-stimulated ATPase which can unwind RNA duplexes (Wagner *et al.*, 1998; Schwer and Gross, 1998).

Another putative ATP-dependent RNA helicase has been isolated and named Prp43p which is probably involved in the disassembly of the spliceosome after the release of mature RNA (Arenas and Abelson, 1998). The mouse homologue of Prp43p has been identified and named mDeah9p (Gee *et al.*, 1998). The precise mode of action of those polypeptides remains unknown.

After release of the snRNPs, the lariat-intron is linearised by a debranching enzyme which specifically cleaves the 2', 5'-phosphodiester bond, then the linearised product is

degraded (Ruskin and Green, 1985 and 1985b; Arenas and Hurwitz, 1987). In yeast cells, *PRP26* codes for a debranching enzyme which is not essential for cell viability (Chapman and Boeke, 1991). Also in yeast cells, mutation in *PRP27* results in accumulation of excised introns (Vijayraghavan, 1989) suggesting that the corresponding gene product may be involved in the debranching of the lariat-intron or the dissociation of the complex.

### 1.4.5 Non-snRNP splicing factors required prior to or during the first catalytic step

Among these factors, the yeast Prp2p, Spp2p, Prp28p proteins and the Prp19-complex are required for the first step of splicing (Kim *et al.*, 1992; Plumpton *et al.*, 1994).

The Prp19 protein is a spliceosomal component, but is not tightly associated with small nuclear RNAs (Tarn *et al.,* 1993; Maddock *et al.*, 1996). It appears to associate with the spliceosome concomitant with, or just after, dissociation of the U4 small nuclear RNA. Prp19p is associated with a protein complex which consists of at least seven other polypeptides (Tarn *et al.*, 1994b). The non-essential yeast *SNT309* gene encodes a polypeptide which is in direct contact with Prp19p and is a component of the Prp19-associated complex (Chen *et al.*, 1998).

Prp2p is required to promote the first step of splicing, leading to the production of lariat-intermediate and exon 1 (Lin *et al.*, 1987). Prp2p interacts only transiently with the spliceosome prior to and during the first step of splicing (King and Beggs, 1990). A mutated form of Prp2p, carrying a mutation in the helicase motif, remains stably associated with the spliceosome blocking splicing before step I (Plumpton *et al.*, 1994), suggesting that some conformational changes are required for this step. Further UV-crosslinking experiments showed that Prp2p contacts the pre-mRNA *in vitro* in an

ATP-independent fashion (Teigelkamp *et al.*, 1994) while ATP is required for the release of Prp2p from the pre-mRNA.

The yeast *SPP2* gene was previously identified as a high-copy suppresser of temperature-sensitive *prp2* mutants (Last *et al.*, 1987). Spp2 is an essential protein required for the first RNA cleavage reaction *in vivo*, and is likely to be released from the spliceosome following ATP hydrolysis by Prp2p (Roy *et al.*, 1995b). The Prp2p and Spp2p proteins are capable of physically interacting with each other.

Prp28p is a member of the DEAD-box family of ATP-dependent putative RNA helicases (see section for description) (Strauss and Guthrie, 1991). Genetic interactions suggested that this polypeptide is associated with the tri-snRNP particle (Strauss and Guthrie, 1991). Its precise function remains obscure, no RNA helicase activity has been demonstrated.

### 1.4.6 Non-snRNP splicing factors required for, or during, the second catalytic step

The yeast splicing factors Prp16p, Prp17p, Prp18p, Slu7p, Ssf1p and Prp22p are required for the second catalytic step (Vijayraghavan and Abelson, 1990; Schwer and Guthrie, 1991; Horowitz and Abelson, 1993; Ansari and Schwer, 1995; Jones *et al.*, 1995).

The earliest defined stage during the second catalytic step is the binding of Prp16p to spliceosomes containing the 5' exon and lariat-intermediate. Prp16p is a protein of the DEAH family of RNA dependent NTPases (Schwer and Guthrie, 1991). Its RNA unwinding activity has been shown *in vitro*, however the lack of sequence specificity does not bring any information about Prp16p's substrate(s) *in vivo* (Wang *et al.*, 1998). Prp16p interacts transiently with the spliceosome and dissociates upon ATP-hydrolysis (Schwer and Guthrie, 1991; Brys and Schwer, 1996). The ATPase activity is essential

for the function *in vitro* and *in vivo* of Prp16p in the splicing reaction, since there is a correlation between reduced ATP hydrolysis and impaired splicing (Schwer and Guthrie, 1992b). ATP-hydrolysis by Prp16p lead to a conformational change in the spliceosome that results in protection of the 3' splice site from oligonucleotide-directed RNase H digestion (Schwer and Guthrie, 1992a). The Prp16p-induced conformational change requires the recruitment of Slu7p, Prp18p and Ssf1p (Ansari and Schwer, 1995; Brys and Schwer, 1996).

Slu7p and Prp18p function during step 2, after Prp16p, in an ATP-independent fashion. Unlike Prp16p, Slu7p remains until the spliceosome is disassembled (Brys and Schwer, 1996). Biochemical and genetic studies implicate Slu7p in 3' splice site selection (Franck *et al.*, 1992; Brys and Schwer, 1996). Slu7p is essential for splicing *in vitro* of precursor RNAs in which the distance between the branchpoint and 3' splice site is >9 nt (Brys and Schwer, 1996). Prp18p is a non-essential protein (Horowitz and Abelson, 1993), only a reduction of the second step of splicing is observed in the absence of Prp18p (Horowitz and Abelson, 1993). Genetic interactions have been reported between *PRP18*, *SLU7* and U5 snRNA (Jones *et al.*, 1995; Frank *et al.*, 1992). Prp18p is dispensable for splicing *in vitro* of precursor RNAs in which the 3' splice site is in close proximity to the branchpoint. It is required only when this distance is >9 nt (Zhang and Schwer, 1997). The need for Prp18p in the second step *in vitro* can be bypassed by excess Slu7p (Zhang and Schwer, 1997).

Ssf1p is required for the second step of splicing, however its function remains unknown (Ansari and Swcher, 1995).

Both *PRP17* and *PRP18* were identified from temperature-sensitive mutants, *prp17-1* and *prp18-1*, that specifically block the second step of splicing (Vijayraghavan *et al.*, 1989, Vijayraghavan and Abelson, 1990). A mutant allele of *PRP17* is synthetic lethal with a loop mutation in U5 snRNA. Prp17p is not essential for splicing as its absence causes only a partial block to the second step of splicing *in vitro* (Horowitz and Abelson, 1993a). Like Prp16p, Prp17p acts before or concomitant with an ATP-

dependent reaction (Horowitz and Abelson 1993b, Jones *et al.*, 1995). It thus acts before Slu7p and Prp18p.

The contribution of Prp22p for the catalysis of step 2 has been recently reported (Schwer and Gross, 1998). Indeed this factor possesses a double function in splicing, it effects the release of mRNAs from the spliceosome in an ATP-dependent fashion and it functions during step 2 in an ATP-independent fashion. Prp22p is required for the splicing of precursor-RNAs only when the distance between the branchsite and the 3' splice site is >21 nt (Schwer and Gross, 1998). It probably acts in concert with Slu7p and Prp18p to position the 3' splice site and the 3'-OH terminal of the 5'exon.

### 1.4.7 Protein motifs in splicing factors

A large variety of protein-protein and protein-RNA interactions occur during spliceosome assembly and the splicing reaction. A number of splicing proteins contain essential characteristic motifs, some of these have been suggested to be directly involved in the protein's function. This is, for example, the case for the polypeptides (Prp2p, Prp5p, Prp16p, Prp28p, Slt22p...) harbouring an NTPase activity and/or an NTP-dependent RNA helicase activity, suggesting that these proteins might be responsible for some RNA conformational changes. Many splicing factors contain an RNA binding motif, a sequence element important for specific, high affinity binding of the protein to the target RNA. Other protein motifs (for example SR repeats) are involved in protein-protein interactions.

#### 1.4.7.1 RNA binding motif

This RNA recognition motif (RRM or RBD) is typically 70-90 amino acids long and contains two short, well conserved elements, RNP1 and RNP2, embedded within a more weakly conserved sequence (reviewed in Birney *et al.*, 1993). Proteins that

contain this motif are capable of binding a variety of RNA conformations including single-strand and stem-loop structures. The RRM is important for efficient interaction between the protein and its target RNA, but multiple sequence elements have the potential to influence RNA-binding specificity (Scherly *et al.*, 1990; Kenan *et al.*, 1991; Boelens *et al.*, 1991; Nelissen *et al.*, 1994).

#### 1.4.7.2 Proteins harboring a DEAD/H box motif

Prp5p, Prp28p and the human UAP56 are members of the DEAD-box family while Prp2p, Prp16p, Prp22p and Prp43p form a subgroup of the DEAH-box proteins (review in Beggs, 1993). Recently the U5 snRNP specific human protein U5-200 kD and its yeast counterpart Slt22p/Brr2p/Rss1p/Snu246p have been shown to belong to another related subgroup (DEXH-box) of RNA helicases. Unexpectedly these proteins harbour two DEXH motifs (Lauber *et al*, 1996). Proteins bearing these motifs belong to a superfamily of putative ATP-dependent RNA helicases. Spliceosome assembly and disassembly involves multiple RNA-RNA interactions that must be continuously formed, rearranged and disrupted to allow each step of splicing to proceed. These RNA conformational rearrangements are most likely mediated by protein factors. In yeast and in mammals, ATP-dependent RNA helicases have been identified as potential candidates for performing this function (reviewed in Will and Lührmann, 1997). The involvement of these proteins at every step in the splicing reaction may account for the requirement of ATP. However, an RNA-unwinding activity has been demonstrated only for few of these proteins (Prp16p, Prp22p, Slt22p) while only an RNA-stimulated ATPase activity as been demonstrated *in vitro* for Prp2p (Kim *et al.*, 1992). It is possible that the failure to detect an RNA helicase activity for these putative helicase proteins may be due to the requirement for additional polypeptide(s) and cofactor(s). This would be analogous to the prototype of the DEAD-box protein family, the eukaryotic transcription factor eIF-4A (Fuller and Lane, 1992). Indeed, eIF-4A requires eIF-4B for helicase activity *in vitro*. The yeast Slt22p/Brr2p/Rss1p was shown to exhibit an RNA-dependent ATPase activity which is particularly stimulated by U2-U6

snRNA duplex, thus providing the first indication of a probable substrate for this RNA helicase (Lin and Rossi, 1996; Noble and Guthrie, 1996; Xu *et al.*, 1996).

#### 1.4.7.3 Zinc-Finger like motif

A number of proteins involved in splicing harbour sequences that relate to the proteins of the zinc-finger family of DNA binding proteins (reviewed in Harrison, 1991). This motif is involved in protein-RNA interactions.

#### 1.4.7.4 RS domain of the SR proteins

In higher eukaryotes, commitment complex formation relies on a complex set of molecular interactions involving members of the SR (serine-arginine-rich) protein superfamily (Staknis and Reed, 1994). These SR proteins are characterised by a domain rich in arginine-serine dipeptides (RS domain) and often one or more RNA-binding domains (RBD) (for review see Fu, 1995; Manley and Tacke, 1996). As a result of their modular organisation, SR proteins can bind to the pre-mRNA and recruit other factors through protein-protein (through their RS domain) or protein-RNA interactions. All SR proteins tested so far are phosphorylated, and studies suggest that both phosphorylation and dephosphorylation play important roles in splicing and splice site selection (reviewed by Mermound *et al.*, 1994). Previous studies investigating the effects of phosphorylation on the interactions between SR proteins have demonstrated that phosphorylation of the RS domain of ASF/SF2 enhances its interaction with the U1-70K protein (Xiao and Manley, 1997). These results support the idea that SR protein interactions are dynamic and regulated by the activity of kinases and phosphatases.

### 1.4.7.5 Isomerases

Some spliceosomal proteins have been proposed to act as isomerases or chaperones, directly promoting protein conformational changes and thereby contributing to the dynamics of the splicing process. Consistent with this hypothesis, recent studies have demonstrated that purified tri-snRNP possess peptidyl-prolyl *cis/trans*-isomerase (PPIase) activity. In particular the tri-snRNPs specific 20 kD protein shares extensive homology with cyclophilin B, a known PPIase (Teigelkamp *et al.*, 1998). Cyclophilins are a highly conserved family of proteins which have been shown to catalyze the *cis-trans* isomerisation of proline residues within peptides (Fisher *et al.*, 1989; Takahasi *et al.*, 1989).

### 1.4.7.6 Gβ-transducin motif

This motif is characterised by a succession of 4-10 repeats of a conserved unit usually ending with the dipeptide Trp-Asp (WD) (reviewed in Simon *et al.*, 1991; Neer *et al.*, 1994; Neer, 1995) The C-terminal region of Prp4p contains these repeated units, which have been shown to be important for its function during splicing (Dalrymple *et al.*, 1989; Hu *et al.*, 1994; Ayadi *et al.*, 1997). Both Prp4p and its human homologue the 60 kD protein, possess seven WD-repeats which can potentially form a seven-bladed propeller structure characteristic of the Gβ subunit of heteromeric G proteins (Wall *et al.*, 1995; Lambright *et al.*, 1996; Sondek *et al.*, 1996). This WD-protein motif is though to function as a protein-protein interactive surface. This is consistent with the fact that several of the known proteins carrying the WD-repeats form multiprotein complexes. This motif was also found in Prp17p (Vijayraghavan *et al.*, 1989; Franck *et al.*, 1992).

### 1.4.8 The U12-type spliceosome

#### 1.4.8.1 Description of the introns spliced by the U12-spliceosome

These introns are found in plants, Drosophila and vertebrates and coexist within genes with introns excised by the U2-spliceosome (for review see Sharp and Burge, 1997). Thus the U12-spliceosome coexists with the U2-spliceosome.
Introns are defined by the 5' and 3' splice sites and by the branchpoint sequences which are shown in Table 3.

**Table 3. Consensus sequences of the introns excised by the U12-type spliceosome.**

| Subclasses of introns | 5' splice site | Branchsite | 3' splice site |
|---|---|---|---|
| **AT-AC** | /AUAUCCUUU | 5'-UUCCUURACYCY-3' | YAC/ |
| **GT-AG** | U/GUAUCCUUU | 5'-UUCCUUAACY-3' | YAG/A |

The structure of the 3' splice site is distinct for these introns. There is typically no polypyrimidine tract and a relative small number of nucleotides (10-16 nts) separating the branchsite from the 3' splice site. Interestingly it has been noticed that this intron 3' end structure is quite similar to that of group II introns (Tarn and Steitz, 1997). This similarity may indicate that the U12-type spliceosome may be more closely related to group II introns than the more common U2-spliceosome.

#### 1.4.8.2 The U12-type splicing machinery

The U12-spliceosome is composed of four low abundance snRNPs U11, U12, U4atac and U6atac (Hall and Padgett, 1996; Tarn and Steitz 1996, 1996b; Kolossova and Padgett, 1997), and it shares the U5 snRNP with the U2-spliceosome.

The secondary structure of the four low-abundance snRNAs U4atac, U6atac, U11 and U12 snRNAs is shown in Figures 9 and 10. These snRNAs and the U5 snRNA are essential for the splicing reaction (Tarn *et al.*, 1996, 1996b; Hall *et al.*, 1996) which involves a two-step trans-esterification mechanism with a lariat-intron-intermediate (Hall and Padgett, 1994; Tarn and Steitz, 1996a).

Analogous RNA-RNA interactions are observed between the U2 and the U12-spliceosomes. Thus the U12 snRNA base-pairs to the branchsite like the U2 snRNA (Hall and Padgett, 1996; Tarn and Steitz, 1996); U11 snRNA was proposed to base-pair to the 5' splice site like the U1 snRNA (Hall and Padgett, 1994). U12 and a fraction of U11 snRNA form a di-snRNP particle (Wassarman and Steitz, 1992), likewise, interactions between a small fraction of U2 and U1 snRNPs have been detected in *Xenopus* (Mattaj *et al.*, 1986). Similarly, U4atac and U6atac are partially base-paired in the U4atac/U6atac snRNP like the U4 and the U6 snRNAs in the U4/U6 snRNP (Tarn and Steitz, 1996b).

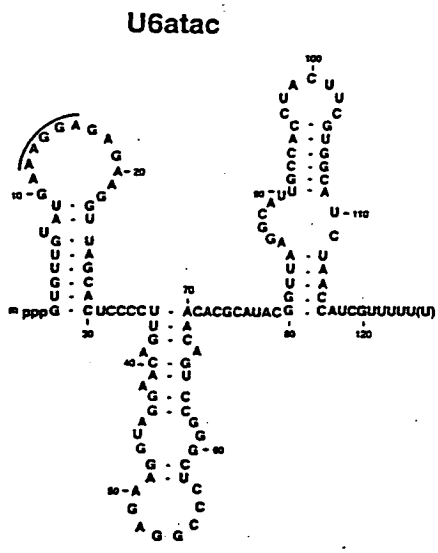
To date, little is known about the proteins involved in the formation of the U12-spliceosome and the splicing reaction. SR proteins (see section 1.4.7 for definition) appear to be required (Tarn and Steitz, 1996a).

## 1.5 This Thesis

This thesis describes the detection of proteins interacting directly with the yeast U6 snRNA. A complex *in vitro* system has been established to study those contacts. Using this system several proteins binding to the U6 snRNA have been detected, and their association with the U6 snRNA in the different snRNP particles and in the spliceosome has been investigated. The identification of these polypeptides in close proximity to the U6 snRNA has been attempted. The effect of mutations or the removal of the last 18 nt of the U6 snRNA molecule on these protein-RNA contacts has been

**Figure 9. Predicted secondary structure of the human U6atac snRNA and U4atac snRNA (Tarn *et al.*, 1996b).**

U6atac snRNA nucleotides potentially involved in base-pairing with the 5' splice site are underlined. Over 32 nucleotides in its central region, U6atac snRNA is ~80% identical to all U6 snRNAs, presumably reflecting a critical function in splicing. The U4atac snRNA is shown base-paired with the U6atac snRNA. Potential RNA-RNA base-pairs are indicated by straight lines. Crosslinking results support the presence of two stem structures (stem I and II) resulting from U4atac/U6atac snRNAs base-pairing interaction. The U4atac and the U6atac snRNAs sequences exhibit only 40% similarity to human U4 snRNA and U6 snRNA respectively. The putative structures of the U4atac snRNA and the U6atac snRNA are more closely related to that of the yeast U4 snRNA and U6 snRNA respectively. Northern blotting analysis revealed that in HeLa cells U6atac snRNA is about three times as abundant as the U4atac snRNA.


U6atac
m pppG CUCCCCU ACACGCAUACG CAUCGUUUUUU(U)

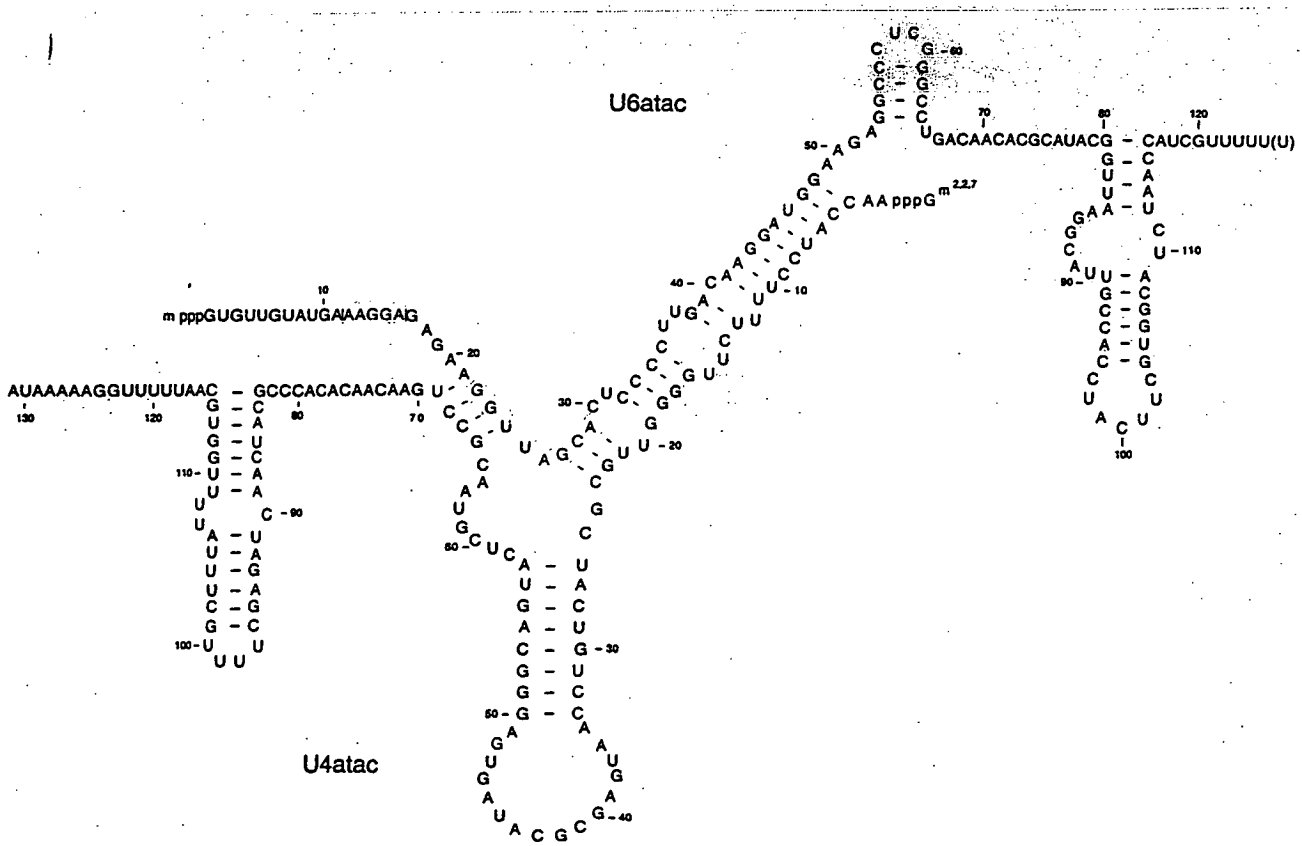
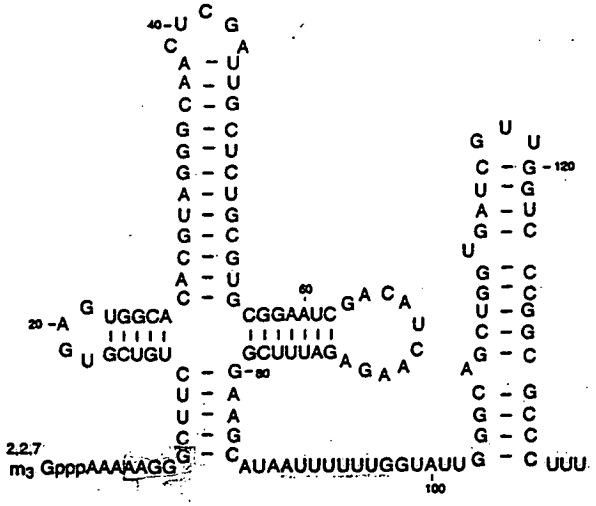

U6atac
m pppGUGUUGUAUGAAAGGAUG
AUAAAAAGGUUUUUAAC GCCCACACAACAAG
GACAACACGCAUACG CAUCGUUUUU(U)
AAppp m2,2,7
U4atac

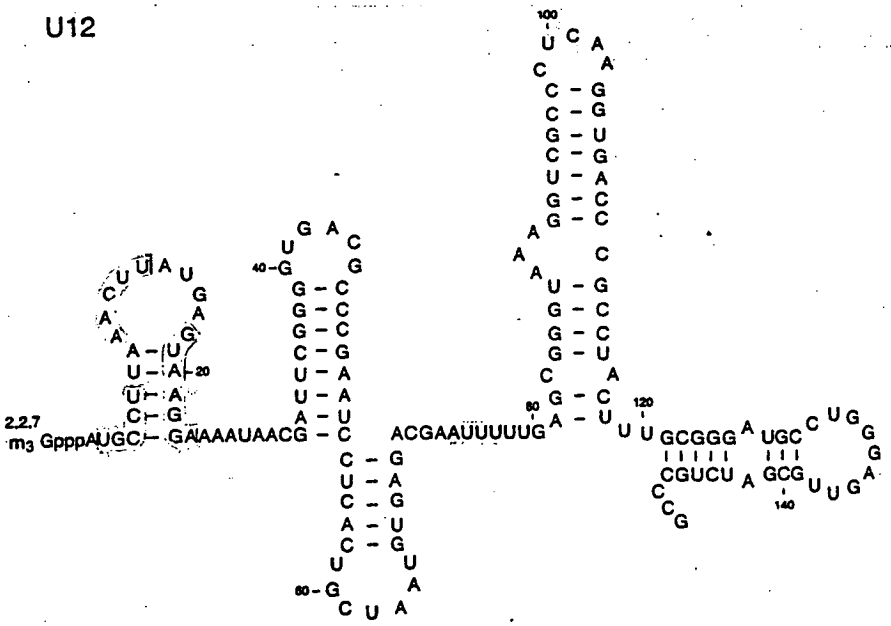
**Figure 10. Potential secondary structure of the low-abundance U11 snRNA and U12 snRNA isolated from HeLa cells (Montza-Wassarman *et al.*, 1992).**

Despite their limited sequence conservation, U11 and U12 snRNAs can fold into secondary structures that are remarkably similar to that of the U1 snRNA and the U2 snRNA respectively. U11 snRNP and U12 snRNP interact together to form a U11/U12 snRNP particle. Their interaction is mediated by proteins yet unknown.

U11


2,2,7
m3 GpppAAAAAGG
AUAAUUUUUUGGUAUU
UGGCA
GCUGU
CGGAAUC
GCUUUAG

U12


2,2,7
m3 GpppAUGC
GAAAAUAACG
ACGAAUUUUUG
GCGGG
UGC
CGUCU
GCG

studied. Also, preliminary experiments have been carried out to map the region of interaction of these proteins to the U6 snRNA molecule.

# CHAPTER TWO

## Materials and Methods

## 2.1 Materials

### 2.1.1 Chemicals and enzymes

Standard laboratory reagents were purchased from Sigma, Fisons, BDH, Serva, Bio-Rad, Boehringer Mannheim.

Phenol: Rathburn Chemicals.

Urea: Fisons.

Acrylamide and N, N'-methylene bisacrylamide stock solutions: Scotlab.

Ultra-Pure Sequagel and Complete Buffer Reagent: National Diagnostics.

Solvents: Scotlab, National Diagnostics.

Agarose: Boehringer Mannheim.

Deoxyribonucleotides: Boehringer Mannheim, United States Biochemical.

Ribonucleotides: Pharmacia.

4-thio$U_pG$ and $U_pG$: Sigma.

Antibiotics: Beecham Research.

BCIP/NBT: Promega.

Growth media reagents: Difco Laboratories, Sigma, Beta Lab., Oxoid.

Xgal: Sigma

Radiochemicals: Amersham.

Restriction endonucleases and DNA modifying enzymes: Gibco BRL, Boehringer Mannheim, New England Biolabs, Pharmacia.

Highly concentrated T4 DNA ligase (2, 000 units/μl): Biolabs.

RNases A and T1: Pharmacia, Boehringer Mannheim.

RNasin (RNase inhibitor): Promega.

T7 and SP6 RNA polymerases: Pharmacia.

T7 DNA polymerase: United States Biochemical.

*Taq* DNA polymerase: Promega, $Vent_R$ DNA polymerase: New England Biolabs and *Pfu* DNA polymerase: Stratagene.

### 2.1.2 Growth media

Except where otherwise indicated, solutions were autoclaved and stored at room temperature.

Antibiotics were added to liquid media immediately prior to use or after autoclaving solid media.

To make agar plates, bacto-agar was added to 2% (w/v) prior to autoclaving.

Given quantities are for one litre volume except when indicated.

#### 2.1.2.1 Bacterial media

Luria-Bertani medium (LB): 10 g bacto tryptone, 5 g yeast extract, 5 g NaCl, adjusted to pH 7.2 with NaOH. For α-complementation test, 40 µl of a stock solution of X-gal (Table 4) and 4 µl of a stock solution of IPTG (Table 4) were spread on the surface of a pre-made agar plate.

YT: 8 g bacto tryptone, 5 g bacto yeast extract, 5 g NaCl.

M9 medium: 6 g $Na_2HPO_4$, 3 g $KH_2PO_4$, 0.5 g NaCl, 1 g $NH_4Cl$, 0.25 g $MgSO_4.7H_2O$, 2 g glucose. After autoclaving, amino acids were added to a final concentration of 40 µg/ml.

#### 2.1.2.2 Antibiotics and α-complementation stock solutions

All the stock solutions indicated below were stored at -20°C.

The antibiotic solutions were diluted one thousand fold in bacterial media.

**Table 4. Antibiotics and α-complementation test stock solutions**

| Chemical | Stock Solution (mg/ml) | Solvent |
|---|---|---|
| **Ampicillin** | 100 | Water |
| **Chloramphenicol** | 30 | 100% Ethanol |
| **Kanamycin** | 50 | Water |
| **IPTG** | 200 | Water |
| **X-gal** | 20 | Dimethylformamide |

2.1.2.3 Yeast media

***YPDA:*** 10 g yeast extract, 20 g bacto peptone, 20 g glucose, 3 mg adenine.

***YPGalA:*** 10 g yeast extract, 20 g bacto peptone, 20 g galactose, 3 mg adenine.

***YMG/CAS:*** 6.7 g yeast nitrogen base w/o amino acids, 20 g casamino acids, 20 g glucose.

***YMGal/CAS:*** 6.7 g yeast nitrogen base w/o amino acids, 20 g casamino acids, 20 g galactose.

***YMM:*** 6.7 g yeast nitrogen base w/o amino acids, 20 g glucose.

***YMGlyLac:*** 6.7 g yeast nitrogen base w/o amino acids, 10 g casamino acids, 2% (v/v) lactate (adjusted to pH 5.7 with KOH), 2% (v/v) glycerol.

For Gal1 promoter induction, galactose was added to a final concentration of 2% (w/v).

***Drop-out medium:*** 6.7 g yeast nitrogen base w/o amino acids, 20 g glucose or galactose, 2 g drop-out powder, adjusted to pH 6 with NaOH.

***Drop-out powder:*** mix 2 g of each amino acid: adenine, alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, tryptophan, uracil, valine, plus 4 g of leucine.

2.1.2.4 Nutrient supplements

**Table 5. Amino acids solutions**

| Constituent | Stock Solution (g per 100 ml) | Final Concentration (mg per L) | Amount of Stock (ml per L of medium) |
|---|---|---|---|
| **Adenine sulfate** | 0.2 | 20 | 10 |
| **Uracil** | 0.2 | 20 | 10 |
| **L-Tryptophan**$^{\phi}$ | 1.0 | 20 | 2 |
| **L-Histidine**$^{\phi}$ | 1.0 | 20 | 2 |
| **L-Methionine**$^{\phi}$ | 1.0 | 20 | 2 |
| **L-Leucine**$^{\phi}$ | 1.0 | 30 | 3 |

$\phi$, Store at $4^0$C.

### 2.1.3 Commonly used buffers

10X PBS$^{\varphi}$: 15 mM $KH_2PO_4$, 43 mM $Na_2HPO_4$, 1.37 M NaCl, 27 mM KCl, pH adjusted to 7.2.

10X TAE: 0.4 M Tris-acetate pH 7.5, 20 mM EDTA.

10X TBE$^{\varphi}$: 0.9 M Tris-borate pH 8.3, 20 mM EDTA

10X TBS$^{\varphi}$: 0.5 M Tris-HCl, 1.5 M NaCl, pH adjusted to 7.5.

10X TE$^{\varphi}$: 0.1 M Tris-HCl, 0.01 M EDTA, pH adjusted to 7.5.

20X SSC: 3 M NaCl, 0.3 M Sodium citrate.

φ, solutions were sterilized by autoclaving.

### 2.1.4 Bacterial strains

**Table 6. Bacterial Strains**

| Strain | Genotype | Source |
|---|---|---|
| **DH5αF'** | F', ϕ80*dlacZ*ΔM15, Δ(*lacZYA-argF*), U169, *deoR*, *rec*A1, *end*A1, *hsd*R17 ($r_K$-, $m_K$+), *supE*44, λ$^-$, *thi*-1, *gyr*A96, *rel*A1 | Gibco BRL |
| **CJ236** | *dut*1, *ung*1, *thi*1, *rel*A1; pCJ105 ($Cm^R$F') | Kunkel et *al.* (1987) |
| **BMH 71-18** | Δ(*lac pro*), *thi*1, *supE*, F'(*lacI*$^q$, *lacZ* ΔM15, *pro*$^+$) | Messing et *al.* (1977) |

### 2.1.5 Yeast strains

**Table 7. Yeast Strains**

| Strain | Genotype | Source |
|---|---|---|
| **S150 2B** | *MATa*, *his*3-1, *leu*2-3, 112, *trp*1-289, *ura*3-52, *GAL*$^+$ | Glaxo |
| **MCY4** | *MATa*, *ade*1-101, *his*3-1, *leu*2-3,112, *trp*1-289, *ura*3-52, *LEU2-GAL1-USS1*, *GAL*$^+$ | This lab. |
| **SC261** | *MATa*, *ura*3-52, *leu*2, *trp*1, *pep*4-3, *prb*1-1132, *prc*1-407 | Gift of I. Kelly Cambridge |
| **Rok8** | *MATa*, *ade*2, *arg*, *leu*2-3, *leu*2-112, *trp*1-289 *ura* 3-52, *snr7::LEU2*, (pROK8; Table 7) | R. O'Keefe (1996) |
| **prp24-a** | *MATa*, *ura*3-52, *lys*2-801, *ade*2-101, *his*3-Δ200, *prp24-4* | U. Vijayraghavan (1986) |
| **prp24-α** | *MATα*, *ura*3-52, *lys*2-801, *ade*2-101, *his*3-Δ*200*, *prp24-4* | U. Vijayraghavan (1986) |
| **DJY84** | *MATα/a*, *ura*3-52/*ura*3-52, *his*3-Δ200/*his*3-Δ*1*, *ADE2*/*ade*2.101$^{oc}$, *LYS2*/*lys* 2-801$^{am}$, *trp*1-Δ*901*/*trp*1-289, *LEU2*/*leu2-3,112*, *GAL*$^+$ | This Lab. |
| **YCLUCK 14** | (Undetermined mating type), *ura*3-52, *his*3, *LEU2*, *LYS2*, *ade*2-101$^{oc}$, *trp*1, *prp24::HIS3*, *URA3-PRP24-CEN* | This work |

| | | |
|---|---|---|
| **YVDP24** | *MATα/a, ura3-52/ura3-52, his3-Δ200/his3-Δ1, ADE2/ade2.101$^{oc}$, LYS2/lys2-801$^{am}$, trp1-Δ901/trp1-289, LEU2/leu2-3,112, prp24::HIS3, GAL$^{+}$* | This Work |
| **AEMY28** | *MATa, trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100, HA:SPB8* | A.Mayes (1998) |
| **AEMY29** | *MATa, trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100,* YER146wΔ*::HIS3*; (pAEM70; Table 9) | A.Mayes (1998) |
| **AEMY30** | *MATa, trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100,* YBL026wΔ*::HIS3*; (pAEM55; Table 9) | A.Mayes (1998) |
| **AEMY31** | *MATa, trp1Δ1, his3-11,-15, ura3-1, leu2-3,-112, ade2-1, can1-100,* YLR438cΔ*::TRP1* (pAEM64; Table 9) | A.Mayes (1998) |
| **AEMY42** | *MATα, trp1Δ1,his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100,* YNL147cΔ*::HIS3*; (pAEM59) | A.Mayes (1998) |
| **AEMY43** | *MATα, trp1Δ1,his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100,* YNL147cΔ*::HIS3*; (pAEM62) | A.Mayes (1998) |
| **AEMY44** | *MATα, trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100,* YDR378cΔ*::HIS3*; (pAEM34, Table 9) | A.Mayes (1998) |
| **AEMY45** | *MATα, trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100,* YDR378cΔ*::HIS3* (pAEM61, Table 9) | A.Mayes (1998) |
| **AEMY46** | *MATa, trp1Δ1, his3Δ200, ura3-1, leu2-3,-112, ade2-1, can1-100,* YBL026wΔ*::HIS3*; (pAEM68; Table 9) | A.Mayes (1998) |
| **LMA4-2A** | *MATa, trp1Δ1, his3-11,15, ura3-1, leu2-3,-112, ade2-1, can1-100,* YJR022wΔ*::TRP1*; (pAEM71; Table 9) | A.Mayes (1998) |

### 2.1.6 Synthetic oligonucleotides

Oligodeoxynucleotides were synthesised at:

- Bioline (London): V538.
- Oswel DNA service (Southampton): T7, G6570, T0314, T0315, P5117, P5118, V4528, S3324, T1850, T1849, T1851, N6088, N6089, S8948, S8949, V5578, S5760, S5761, W3464, W3462, W3463.
- Perkin Elmer Ltd. service (Warrington): JMB10, JMB11, Oligo1, Oligo2, Oligo3, SRU2.
- Genosys Biotechnologies Ltd. service (Cambridgeshire): BFUO, FU6, T7b, US383, U359, UB59, UB38, BU80, UFO, U559, U380, U383.
- Oligos 483A, and 485A were a gift from D. Brow. Oligo Rok8L and Rok8R were a gift from A. Newman.

**Table 8. Oligodeoxynucleotides**

| Oligo | Description | Sequence (5'-3') |
|---|---|---|
| **486A** | Complentary to yeast U1 snRNA sequence. | CTTAAGGTAAGTAT |
| **G1803** | Complementary to yeast U2 snRNA sequence. | CTACACTTGATCTAAGCCAAA AGGC |
| **SRU2** | U2 snRNA knockout oligo. Complementary to yeast U2 snRNA sequence 29-43. | CAGATACTACACTTG |
| **483A** | Complementary to yeast U4 snRNA sequence 1-14. | CCGTGCATAAGGAT |
| **T0314** | Complementary to yeast U4 snRNA sequence 72-92. | GAGACGGTCTGGTTTATAATT |

| | | |
|---|---|---|
| **T0315** | Complementary to yeast U4 snRNA sequence 65-84. | CTGGTTTATAATTAAATTTC |
| **485A** | Complementary to yeast U5 snRNA sequence 156-169. | AATATGGCAAGCCC |
| **Rok8L** | U5 snRNA knockout oligo | TTTCTCCCATGTTCGTTATA |
| **Rok8R** | U5 snRNA knockout oligo | CGGATGGTTCTGAAGAACCAT GTT |
| **Lory1** | T7 promoter sequence followed by six extra nucleotides (which improve the transcription efficiency) and the U6 snRNA gene sequence 1-11. | CGTAATACGACTCACTATAGG GAGAGTTCGCGAAGTAACCC |
| **P0052** | Complementary to the *PRP24* ORF sequence 270-304 followed by sequence 338-321 of the *HIS3* gene. | GGCGCCATTGAAGTCGCTAAC GAGAAGTATTTTCCCTCTTGG CCTCCTCTAG |
| **P0052** | Contains the *PRP24* ORF sequence 64-92 followed by sequence 1330-1313 of the *HIS3* gene. | GCCGCAGGATTAACTTCTAAG AAGGCGAACGAAGCTCGTTC AGAATGACACG |
| **S5760** | T7 transcription promoter sequence followed by the yeast U6 snRNA gene sequence 1-7, an extra G was added at the 5' end of the wild-type U6 snRNA gene sequence. | CGTAATACGACTCACTATAGG TTCGCG |
| **S5761** | Complementary to the 3'end of the yeast U6 snRNA sequence 85-112. | AAAACGAAATAAATCTCTT TGTAAAAC |
| **Taq6A** | Complementary to yeast U6 snRNA sequence 43-56. | TCATCTCTGTATTG |
| **T1849** | Complementary to yeast U6 snRNA sequence 39-56 with a mismatch at position C48. | TCATCTCTTTATTGTTTC |

| | | |
|---|---|---|
| **T1850** | Complementary to yeast U6 snRNA sequence 52-72 with a mismatch at position G60. | GGGAACTGGTGATCATC |
| **T1851** | Complementary to yeast U6 snRNA sequence 56-68 with a mismatch at position A59. | GGAACTGCAGATCATC |
| **V4528 and S3324** | Yeast U6 snRNA knockout oligo. Complementary to yeast U6 snRNA sequence 28-54. | ATCTCTGTATTGTTTCAAATTG ACCAA |
| **V5578** | Complementary to yeast U6 snRNA sequence 81-98. | CTCTTTGTAAAACGGTTC |
| **W3462** | Complementary to yeast U6 snRNA sequence 34-74. | ATGCAGGGGAACTGCTGATCA TCTCTGTATTGTTTCAAATT |
| **W3363** | Complementary to the T7 transcription promoter sequence followed by the yeast U6 snRNA gene sequence 55-74. | GCGCTAATACGACTCACTATA TGATCAGCAGTTCCCCTGCAT |
| **W3364** | Complementary to yeast U6 snRNA sequence 33-53. | TCTCTGTATTGTTTCAAATTG |
| **T7b** | T7 transcription promoter sequence. | GCGCTAATACGACTCACTATA |
| **JMB10** | Complementary to *PRP24* ORF downstream sequence. | CGCCTTTCATGAACAG |
| **JMB11** | Complementary to *PRP24* ORF sequence 966-983. | GGTGTCACCAAGCTTAATC |
| **N6088** | Prp24p coding sequence 1320-1332, the A of the TAG stop codon was changed by a C introducing a *Bam*HI restriction site, eight extra nucleotides were added at the 3' end. | CCCGGATCCTCACCTAGAAAC |

| | | |
|---|---|---|
| **N6089** | Complementary to *PRP24* ORF sequence 437-451. | GTATACGGCTTCCCAG |
| **Oligo1** | Prp24p encoding sequence 83-98. | GCTTCGTTCGCCTTC |
| **Oligo2** | Prp24p encoding sequence 210-226. | CGCAATCCGCAACATC |
| **Oligo3** | Prp24p encoding sequence 628-675. | GCGCCCTCTAGTGTG |
| **P5117** | Complementary to Prp24p encoding sequence 1050-1065. | GGCGCCATTGAAGTC |
| **P5118** | Complementary to Prp24p encoding sequence 83-98. | GAAGGCGAACGAAGC |
| **S8948** | Complementary to the 5' end of Prp24p encoding sequence 1-20, contains a *Bam*HI site. | CCCGGATCCATGGAGTATGGA CATCACGC |
| **S8949** | Complementary to the 3' end of Prp24p encoding sequence 1320-1333, contains an *Xba*I site. | CCACTCTAGAGTCTACTCACC TAGAAAC |
| **M5392** | Stop codon linker. | TCGATTAACTAACTAG |
| **M5393** | Stop codon linker. | TCGACTAGTTAGTTAA |
| **N4804** | Flag linker. | TCGACCCGGGGGATCCAGACT ACAAGGACGACGATGACAAG CTT |
| **N4805** | Flag linker. | TCGAAAGCTTGTCATCGTCGT CCTTGTAGTCTGGATCCCCCG GG |

### 2.1.7 Plasmid vectors and constructs

**Table 9. Plasmid vectors and constructs**

| Plasmid | Features | Reference |
|---|---|---|
| **pBluescript KS$^{+/-}$** | Phagemid cloning vector: Multiple cloning site flanked by T3 and T7 transcription promoters, f1 origin, *lacZα*, *Amp*$^R$. | Stratagene. |
| **pATH3** | *E. coli* expression vector: Used to produce an incomplete *trpE* fusion protein. Expression of the fusion protein controlled by the indoleacrylic acid-inducible *trp* promoter. Multiple cloning site; *Amp*$^R$. | Koerner et *al.* (1990) |
| **pGEX-2T** | *E. coli* expression vector: Used to produce a glutathione S-transferase (GST) fusion protein. Expression of the fusion protein controlled by the IPTG-inducible *tac* promoter. A designed thrombin protease recognition site can be used to cleave the protein of interest from the fusion product. Multiple cloning site; *Amp*$^R$, *lacI*$^q$. | Pharmacia. |
| **pUR288** | *E. coli* expression vector: Used to produce a β-galactosidase fusion protein. Expression of the fusion protein is controlled by the IPTG-inducible *lac* promoter. Multiple cloning site; *Amp*$^R$. | Rüther et *al.* (1983) |
| **pQE30** | *E. coli* expression vector: Used to produce a 6X Histidine fusion protein. Expression of the fusion protein is controlled by the IPTG-inducible *lac* promoter. Multiple cloning site; *Amp*$^R$. | Qiagen. |
| **pBM125** | Yeast-*E. coli* shuttle vector: Gal1-Gal10 promoter sequence, *CEN4*, *ARS1*, *URA3*, *tet*$^R$, *Amp*$^R$. | Johnston et *al.* (1984) |
| **YCp LAC22** | Yeast-*E. coli* shuttle vector: Multiple cloning site, *lacZ*, *Amp*$^R$, *CEN4*, *ARS1*, *TRP1*. | Gietz *et al.* (1988) |

| | | |
|---|---|---|
| **pATP24** | Modified pATH3: Contains Prp24p encoding sequence amplified by PCR from yeast genomic DNA with primers S8948 and S8949. The PCR product was *Bam*HI/*Xba*I digested, gel purified and cloned in *Bam*HI/*Xba*I-opened pATH3 vector. | This work. |
| **pBXPR24** | Modified pBluescript: The *Xba*I-Prp24p containing ORF fragment isolated from plasmid YCpXba was cloned in *Xba*I-opened pBluescript. | This Work. |
| **pGEX24** | Modified pGEX-2T: Contains Prp24p coding sequence in frame with the GST gene. The Prp24p gene was isolated as a *Bam*HI/*Sal*I restriction fragment from pATP24 plasmid, and cloned in 3 steps in *Bam*HI/*Sma*I opened pGEX-2T vector: First, the *Bam*HI compatible Vector/insert ends were ligated together. Then the *Sal*I-3' recessed end of the insert was filled-in by the klenow enzyme. Finally, the blunted 3' end of the Prp24p gene was ligated to the vectors' *Sma*I end. | This work. |
| **pGAP24** | Modified pUR288: contains Prp24p coding sequence in frame with the β-galactosidase gene. *PRP24* ORF was gel isolated from *Bam*HI/*Xba*I-digested pGEX24 plasmid and cloned in *Bam*HI/*Xba*I opened pUR288 vector. | This work. |
| **pIP24** | Modified pQE30: contains *PRP24* ORF cloned in frame with the 6X histidine encoding sequence. *PRP24* ORF was gel isolated from *Bam*HI/*Sal*I-digested pATP24 plasmid and cloned in *Bam*HI/*Sal*I opened pQE30 vector. | This work |
| **YCpXba** | Modified Ycp50: Contains a 3.5 Kb *Xba*I fragment harboring the Prp24p coding sequence with upstream and downstream sequences. | Shannon *et al.* (1991) |
| **YCpL24** | Modified YCpL22: Contains the *Xba*I fragment of YCpXba cloned into *Xba*I-opened YCpLac22. | This work. |

| | | |
|---|---|---|
| **YCR24** | Modified YCpL24: This construct was *ClaI*/*Bam*HI digested and the *ClaI*/*Bam*HI fragment replaced by a PCR product obtained using oligos N6088 and N6089. This PCR product was *ClaI*/*Bam*HI digested and contains the 3' end of the *PRP24* ORF. The resulting YCR24 plasmid possesses the Prp24p coding sequence but no stop codon. | This work. |
| **YRFL24** | Modified YCR24: The Flag-Stop cassette from pBFAST was isolated by *Bam*HI digestion and inserted into YCR24 linearised with *Bam*HI. The Flag epitope coding sequence was cloned in frame at the 3' end of the *PRP24* ORF. | This work. |
| **pBM24** | Modified pBM125: *PRP24* ORF was isolated from pATP24 *Bam*HI-*Sal*I digested and inserted into *Bam*HI-*Sal*I cut pBM125 plasmid. | This work. |
| **pBSTP** | Contains stop codons in 3 different ORFs: Stop linkers M5392 and M5393 were phosphorylated, annealed and cloned into *Sal*I cut pBluescript. | This Work. |
| **pBFAST** | Contains the Flag epitope coding sequence followed by stop codons in all three frames: The FLAG linkers N4804 and N4805 were phosphorylated, annealed and inserted into pBSTP linearized with *Sal*I. | This Work. |
| **pACTII** | Yeast-*E. coli* shuttle and expression vector used for the two-hybrid system: The gene of interest can be cloned in frame with the sequence coding for the Gal4 DNA binding domain and the HA-epitope. The expression of the gene is under the control of the constitutive *ADH1* yeast promoter. Multiple cloning site, 2μ origin, *TRP1*, f1$^+$ origin, *Amp*$^R$. | Clontech |
| **pAS2ΔΔ** | Yeast-*E. coli* shuttle and expression vector used for the two-hybrid system: The expression of the inserted gene is under the control of the constitutive *ADH1* yeast promoter. The protein of interest is produced as a fusion protein with the Gal4 DNA binding domain. Multiple cloning site, 2μ origin, *TRP1*, f1$^+$ origin, *Amp*$^R$. | Fromont-Racine *et al.* (1997) |

| | | |
|---|---|---|
| **pBTM116** | Yeast-*E. coli* shuttle and expression vector used for the two-hybrid system: Use to produce a LexA DNA binding domain fusion protein. The expression of the gene is under the control of the constitutive *ADH1* yeast promoter. Multiple cloning site, 2μ origin, *TRP1*, $f1^+$ origin, $Amp^R$. | Gift from A. Brown. (Glaxo) |
| **YCpIF16** | Yeast expression vector: Use to overproduce an HA epitope-tagged protein. This vector contains a sequence encoding an HA epitope upstream from the multiple cloning site. The expression of the gene is under the control of the *GAL1* promoter. *CEN4*, *TRP1*. | Foreman and Davis. (1994) |
| **pAEM13** | Modified pAS2ΔΔ: The Uss1p coding sequence was cloned in frame with the Gal4 DNA binding domain. | A. Mayes. (This Lab.) |
| **pAEM34** | Modified pBTM116: The YDR378c ORF was cloned in frame with the sequence coding for the LexA DNA binding domain. | A. Mayes. (This Lab.) |
| **pAEM55** | Modified pACTII: YBL026w ORF was cloned in frame with sequence coding for the Gal4 DNA binding domain and the HA-epitope. | A. Mayes. (This Lab.) |
| **pAEM59** | Modified pBTM116: The YNL147c ORF was cloned in frame with the sequence coding for the LexA DNA binding domain. | A. Mayes. (This Lab.) |
| **pAEM61** | Modified YCpIF16: The YDR378c ORF was cloned in frame with the sequence coding for the HA epitope. | A. Mayes. (This Lab.) |
| **pAEM62** | Modified YCpIF16: The YNL147c ORF was cloned in frame with the sequence coding for the HA epitope. | A. Mayes. (This Lab.) |
| **pAEM64** | Modified pBM125: Contains the HA-YLR438c ORF coding for the HA-Uss2 fusion protein. | A. Mayes. (This Lab.) |

| | | |
|---|---|---|
| **pAEM68** | Modified pBM125: Harbours the YBL026w-HA ORF encoding Uss6p fused at its C-terminus with the HA epitope. | A. Mayes. (This Lab.) |
| **pAEM70** | Modified pACTII: YER146w ORF was cloned in frame with sequence coding for the Gal4 DNA binding domain and the HA-epitope. | A. Mayes. (This Lab.) |
| **pAEM71** | Modified pACTII: YJR022w ORF was cloned in frame with the sequence coding for the Gal4 DNA binding domain and the HA-epitope. | A. Mayes. (This Lab.) |
| **pGPRP24** | Modified pAS2ΔΔ: The *PRP24* ORF was cloned in frame with the Gal4 DNA binding domain. | A. Colley. (This Lab.) |
| **pUC12U6** | Contains the yeast *SNR6* gene cloned in pUC12 vector. This gene possesses the T7 promoter sequence and an extra G nucleotide at its 5' end. This plasmid was linearized with *Dra*I for run-off transcription. | I. Kelly. (Cambridge) |
| **pU6ribo** | Modified pUC12U6: A hairpin ribozyme was cloned 3' to the U6 gene sequence. This plasmid was linearized with *Eco*RI for run-off transcription. | I. Kelly. (Cambridge) |
| **pUCU6f1** | Modified pBluescript $KS^+$: pUC12U6 plasmid was *Bgl*I digested and the fragment containing the T7-SNR6 gene was cloned in *Bgl*I-opened pBluescript $KS^+$ vector. | I. Kelly. (Cambridge) |
| **pU6C48A** | Modified pUC6f1: The *SNR6* wild-type gene sequence was subjected to site directed mutagenesis using DNA oligo T1849. C48 nucleotide was mutated to A. | This work. |
| **pU6A51U** | Modified pUC6f1: The *SNR6* wild-type gene sequence was subjected to site directed mutagenesis. A51 nucleotide was mutated to T. | I. Kelly. (Cambridge) |

| | | |
|---|---|---|
| **pU6A59U** | Modified pUC6f1: The *SNR6* wild-type gene sequence was subjected to site directed mutagenesis using DNA oligo T1851. A59 nucleotide was mutated to T. | This work. |
| **pU6G60C** | Modified pUC6f1: The *SNR6* wild-type gene sequence was subjected to site directed mutagenesis using DNA oligo T1850. G60 nucleotide was mutated to C. | This work |
| **p283** | Modified pGEM1 vector: Contains an *Alu*I fragment of the yeast *ACT1* gene inserted at the *Sma*I site in the T7 orientation. The plasmid was linearized with *Bam*HI for run-off transcription. | O'Keefe *et al.* (1996) |
| **pROK8** | Harbours a modified *SNR7* gene coding for the U5 snRNA containing a 30 nucleotide-insertion required for U5 snRNA depletion. | O'Keefe *et al.* (1996) |
| **pSPrp51A** | Contains the RP51A gene. SP6 run-off transcripts were generated with a 100 nt long poly(A) tail when linearized with *Bam*HI; this tail was missing when linearized with *Eco*RI. | Pikielny et *al.* (1986) |
| **pBM-PRP2$^{LAT}$** | Modified pBM125: Contains the dominant negative mutant Prp2$^{LAT}$p encoding gene under the control of the *GAL1* inducible promoter. | This Lab. |

### 2.1.8 Antisera

**Table 10. Description of antibodies**

| Antibody | Description | Origin |
|---|---|---|
| **Anti-8.6 (666)** | Rabbit polyclonal antibodies: Raised against a 35 amino acid synthetic peptide of the N-terminal region of Prp8p. | This lab. |
| **Anti-PAB (170)** | Rabbit polyclonal antibodies: Raised against the yeast poly(A)-binding protein. | Whittaker *et al.* (1991) |
| **Anti-Prp2 (140)** | Rabbit polyclonal antibodies: Raised against an internal region of Prp2p as β-galactosidase fusion protein. | Teigelkamp *et al.* (1994) |
| **Anti-Uss1 (756 and 763)** | Rabbit polyclonal antibodies: Raised against the entire Uss1p as a β-galactosidase fusion protein. | This lab. |
| **Anti-Prp24 (1015 and 1016)** | Rabbit polyclonal antibodies: Raised against the entire Prp24p as a β-galactosidase fusion protein. | This work. |
| **Anti-Rabbit IgG** | Alkaline phosphatase conjugated anti-rabbit IgG (Fc) antibodies. | Promega. |
| **Anti-HA** | Monoclonal antibodies raised against the haemagglutinin (HA) peptide. | Gift from D. Xu (University of Toronto) |

## 2.2 General Methods

### 2.2.1 Deionization of solutions

Formamide and Acrylamide/Urea solutions were deionized by mixing with 0.1 volume of analytical mixed-bed resin (20-50 mesh; Bio-Rad Laboratories), stirring for 30 min. The resin was removed by filtration through Whatman No.1 filter paper.

### 2.2.2 Autoradiography

[$^{32}$P]-labelled nucleic acids in gels or on membrane were detected by exposure to Cronex X-ray film (Du Pont) at -70°C in a light-proof cassette with a calcium tungstate intensifying screen. For the detection of [$^{32}$P.RNA-protein] UV-crosslinking signals, a hypersensitive film (Biomax MS, Kodak) was used in combination with a Biomax MS intensifying screen (Kodak).
[$^{35}$S]-labelled DNA was also detected by exposure to Cronex X-ray film, at room temperature and without any intensifying screen.

### 2.2.3 Cerenkov counting

Cerenkov counting of [$^{32}$P] was performed in a Tri-Carb 2 100TR Packard (Canberra compagny) scintillation counter.

## 2.3 Microbiological Methods

### 2.3.1 Propagation and storage of strains

2.3.1.1 *E. coli* strains

Strains were routinely grown at 37°C, in rich LB medium (Section 2.1.2.1) or in minimal M9 medium supplemented with nutrients (Table 5) if necessary. To maintain the plasmid DNA selection in bacteria, antibiotics were added to growth media (Table 4).

2.3.1.2 Yeast strains

Strains were routinely grown at 30°C (or different temperature according to strain thermosensitivity) in complete YPDA or YPGalA media (Section 2.1.2.3). To maintain plasmid DNA selection transformed yeasts were grown in selective media supplemented with amino acids (Table 5) according to strain genotype (Table 7), or in drop-out medium lacking the relevant amino acids.

Both yeast and bacterial strains were preserved for short periods at 4°C on plates. For long term storage: 400 µl of yeast or *E. coli* liquid culture were transferred to a 2 ml cryovial containing 600 µl of 100% sterile glycerol. The contents were gently mixed and rapidly frozen in dry ice. The vial was kept at -70°C for an indefinite period.

### 2.3.2 Transformation of *E. coli*

2.3.2.1 Preparation of competent cells

Competent cells were prepared according to a modified method of Hanahan (1983): a culture was grown overnight at 37°C to stationary phase in 5 ml LB-medium (Section 2.1.2.1). The following day 1 ml of this culture was used to inoculate 100 ml

LB-medium. Cells were grown at 37°C for 2-3 h (until $OD_{550}$= 0.5). The culture was transferred to sterile JLA 10-500 centrifugation bottles (Beckman, Avanti™ J 25 centrifuge), cooled on ice for 10 min, then sedimented by centrifugation for 10 min at 4000 rpm (4°C). The supernate was discarded, the pellet resuspended in 17 ml of ice-cold Tfb-1 solution and kept on ice for at least 20 min. The cells were sedimented by centrifugation for 10 min at 4000 rpm (4°C) and resuspended very gently in 8 ml of ice-cold Tfb-2 solution. Competent cells were aliquoted in Eppendorfs tubes, snap-frozen in dry ice and stored for months at -70°C.

- Tfb-1 solution: 30 mM KOAc, 50 mM $MnCl_2$, 100 mM KCl, 10 mM $CaCl_2$, 15% (v/v) glycerol.

- Tbf-2 solution: 10 mM Na-MOPS pH 7.0, 75 mM $CaCl_2$, 10 mM KCl, 15% (v/v) glycerol.

Both solutions were sterilized by filtration through a sterile acrodisc filter (0.45 µm) (Gelman Sciences).

#### 2.3.2.2 Transformation of competent cells

Frozen competent cells were thawed on ice, 50-200 µl aliquots of cells were gently mixed with transforming DNA. The mixture was incubated on ice for 30 min then heat shocked at 42°C for 45 sec and placed on ice for 2 min. 1 ml of LB-medium was then added and the cells incubated for 1 h at 37°C on a rotating wheel. Aliquots (50-200 µl) of diluted or undiluted cells were plated onto selective agar medium. To assay for the lac-complementation, plates were supplemented with X-gal/IPTG (section 2.1.2.1) and incubated overnight at 37°C.

### 2.3.3 Transformation of yeast

Yeast transformation was carried out according to the method of Gietz *et al.* (1992).

Yeast cells to be transformed were inoculated into 10 ml liquid YPDA medium (section 2.1.2.3) and grown overnight to 1-2 x $10^7$ cells/ml. The following day the cells were diluted to 2 x $10^6$ cells/ml in fresh liquid YPDA medium and regrown to 1 x $10^7$ cells/ml. Cells were then harvested by centrifugation at 3000 rpm for 5 min (Mistral 1000 centrifuge) in sterile Falcon tubes (Greiner). The supernate was discarded and the cell pellet washed in 30 ml sterile water. The cells were pelleted by centrifugation, washed twice in 1 ml water and transferred into a sterile 1.5 ml microfuge tube. The cells were washed in 1 ml 1X TE/LiAc (made fresh from 10X autoclaved TE and LiAc stock solutions) and resuspended at 2 x $10^9$ cells/ml in 1X TE/LiAc. In a 1.5 ml microfuge tube, 50 µl of the yeast cell suspension was mixed with 1 µg transforming DNA, 50 µg of single-stranded salmon sperm carrier DNA and 300 µl of sterile 40% PEG/TE/LiAc solution (made fresh from 50% PEG 3500, 10X TE and 10X LiAc stocks). The mixture was homogenized by inversion of the tube and incubated at 30°C for 30 min on a rotating wheel. Cells were heat-shocked at 42°C for 15 min then spun down in a microfuge for 5 sec. The supernate was discarded and the cells gently resuspended in 1 ml of 1X TE (made fresh from a 10X stock). The cell suspension was diluted appropriately and plated onto selective medium. Plates were incubated at the appropriate temperature for 2-3 days until the appearance of colonies. Cells from a yeast colony were restreaked onto selective plates for colony isolation.

- 10X TE: 0.1 M Tris-HCl, 0.01 M EDTA, pH adjusted to 7.5.
- 10X LiAc: 1 M LiAc pH adjusted to 7.5 with diluted acetic acid.

### 2.3.4 Sporulation of yeast

#### 2.3.4.1 Growth of diploid cells in liquid medium

YPDA medium (10 ml) (section 2.1.2.3) was inoculated with the diploid yeast strain to be sporulated and the culture grown at 30°C to an optical density at 600 nm of 2.5-3.0. 1 ml of cells was transferred to a sterile Eppendorf, washed once with sterile water and resuspended in 1 ml of liquid sporulation medium supplemented with nutrients (Table 5) required for the growth of the diploid strain. The cells were grown in sporulation medium for at least 3 days at 30°C to induce tetrad formation. Tetrad formation was observed under the microscope.

- Liquid sporulation medium: 10 g/l potassium acetate, 1g/l bacto yeast extract, 0.5 g glucose, relevant amino acid (Table 5).

#### 2.3.4.2 Growth of diploid in solid medium

Diploids to be sporulated were grown on YPDA or selective plates for 3-4 days. A small dab of cells was collected from the plate and washed in 1 ml sterile water. Cells were resuspended in 100 µl of sterile water and spread onto a sporulation plate (Sporulation medium supplemented with 2% (w/v) agar). The plate was left at 30°C for at least 4 days or at lower temperature which sometime increases the sporulation efficiency. Tetrads were visualized by microscopy.

#### 2.4.3.3 Tetrad dissection

Following growth in sporulation medium the culture was transferred to an Eppendorf, the cells pelleted by centrifugation, washed twice with sterile water, then resuspended in 5 ml sterile water. 30 µl of cells were diluted in 270 µl of sterile distilled water and incubated with 20 µl of β-Glucuronidase/Arylsulfatase (Boehringer

Mannheim) for 30-60 min at room temperature. Cell wall digestion was analysed microscopically during this incubation period. Tetrads were dissected on a thin YPDA plate using a Singer MSM system micro-manipulator.

## 2.4 Nucleic Acid Methods

### 2.4.1 Spectrophotometric determination of DNA and RNA concentrations

Nucleic acid concentrations were determined by measuring the absorption of diluted solutions at 260 nm using a Cecil CE 2040 spectrophotometer and a quartz cuvette (Cecil instruments).

For double-stranded DNA an $OD_{260}$ value of 1 represents a DNA concentration of 50 μg/ml, and 40 μg/ml of single-stranded DNA or RNA.

For single-stranded RNA an $OD_{260}$ value of 1 represents a RNA concentration of 37 μg/ml.

### 2.4.2 Precipitation of nucleic acids

Nucleic acids were precipitated with ethanol in the presence of salts. The salt concentration was adjusted by adding 0.1 volume of 3 M NaOAc (pH 5.3) or by adding 0.5 volume of 7.5 M $NH_4OAc$ solutions plus 2.5 volume of 100% ice-cold ethanol and the mixture inverted several times. Nucleic acids were precipitated at -20°C for 30 min or at -70°C for 15 min, followed by centrifugation at 4°C for 15 min. The nucleic acids pellet was washed with 70% ethanol, dried under vacuum and resuspended in the appropriate volume of 1X TE buffer (section 2.1.3).

### 2.4.3 Deoxyribonucleic acid methods

#### 2.4.3.1 Plasmid DNA preparation

***Small scale preparation***

The following protocol is based on the alkaline lysis method of Zhou *et al.* (1990).

LB-medium (5 ml, containing selective antibiotic, section 2.1.2.1) was inoculated with a single colony of the plasmid-bearing *E. coli* strain and incubated overnight with continuous shaking at 37°C. A 2 ml aliquot of the culture was centrifuged for 2 min, the supernate discarded and the cell pellet resuspended by vortexing in 300 µl of TENS solution. 150 µl of 3 M NaOAc (pH 5.2) was added to the cell suspension and the mixture vortexed 2-5 sec. The cell debris and chromosomal DNA were pelleted by centrifugation at high speed for 2 min in a table centrifuge. The pellet was removed from the supernate with a sterile toothpick. Residual protein debris was removed from the supernate by phenol:chloroform extraction. To the aqueous phase, 900 µl of precooled 100% ethanol was added, and the nucleic acids pelleted by centrifugation at high speed for 5 min. The pellet was washed with 70% ethanol, dried and resuspended in 30 µl of 1X TE/RNase A. The resuspended DNA was incubated at 37°C for 30 min then frozen.

- 1X TE/RNase A: Prepared from 10X TE stock solution (section 2.1.3), contains RNase A at a final concentration of 20 µg/ml.

- TENS: 1X TE, 0.1 M NaOH, 0.5% (w/v) SDS.

***Medium scale preparation***

LB-medium (50 ml, containing selective antibiotic, section 2.1.2.1) was inoculated using 1 ml of a saturated plasmid-bearing *E. coli* strain and incubated at 37°C with shaking overnight. The culture was centrifuged for 10 min at 4000 rpm in a

Beckman centrifuge (Beckman JLA 10-500 rotor). The supernate was discarded and the cell pellet resuspended in 4 ml of LETR solution. This cell suspension was transferred to a 50 ml polycarbonate tube and incubated at room temperature for at least 30 min with occasional mixing by inversion. 80 µl of 10% (v/v) Triton X-100 in TE was added, the mixture homogenized by inversion and incubated at room temperature for 10-30 min. The cell debris and chromosomal DNA were pelleted by centrifugation for 30 min at 12000 rpm in a Beckman centrifuge (Beckman JA-25 rotor). The DNA-containing supernate was transferred into a new polycarbonate tube and 300 µl of 3 M NaCl was added. Proteins were extracted by addition of 4 ml phenol/chloroform solution. Aqueous and organic phases were separated by centrifugation for 20 min at 10000 rpm (4°C). The organic phase was discarded, 1.5 ml of chloroform was added and the mixture vigorously shaken followed by centrifugation for 5 min at 10000 rpm. The aqueous phase was transferred to a fresh tube and 0.5 volume of isopropanol was added. The mixture was incubated at room temperature for 10 min, and the plasmid DNA pelleted by centrifugation for 30 min at 11000 rpm (4°C). The supernate was discarded and the pellet washed with 75% ethanol, dried and dissolved in 0.5-1 ml TE. The DNA solution was transferred in a sterile Eppendorf tube and stored at 4°C for an indefinite time.

- LETR solution: 2 g/l lysozyme, 100 mM EDTA, 50 mM Tris-Cl pH 8.0, 0.1 mg/ml RNaseA, 10% (v/v) glycerol.

***Large scale preparation***

LB-medium (500 ml, containing selective antibiotic, section 2.1.2.1) was inoculated using 5 ml of a saturated plasmid-bearing *E. coli* strain and incubated at 37°C with shaking overnight. The culture was centrifuged for 10 min at 4000 rpm in a Beckman centrifuge (Beckman JLA 10-500 rotor). The supernate was discarded and the cell pellet resuspended in 7.5 ml of [25% (w/v) sucrose, 50 mM Tris-Cl pH 7.5] solution. This cell suspension was transferred in a 50 ml polycarbonate tube and left on ice for 5 min then 2.5 ml of fresh 10 mg/ml lysozyme solution in water was added and incubated on ice for 5 min. 3 ml of 0.25 M EDTA pH 8.0 were added and the mixture

kept on ice for 5 min, then 12 ml of lysis buffer was added and the mix incubated for an extra 30 min on ice. The cell debris and chromosomal DNA were pelleted by centrifugation for 30 min at 16000 rpm (4°C) in a Beckman centrifuge (Beckman JA-25 rotor). The DNA-containing supernate was transferred into a Falcon tube and the volume adjusted to 30 ml with water. 30 g of CsCl was added and the salt carefully dissolved by inversion of the tube until complete dissolution. 1.5 ml of a 10 mg/ml ethidium bromide solution was added and the mix was separated in two 18 ml Quickseal tubes (Beckman). The tubes' weight were precisely equilibrated (between each others): A maximal weight difference of 10 mg was allowed. The tubes were spun for at least 17 h at 45000 rpm (room temperature) in a Beckman XL-100 Ultracentrifuge (Rotor Ti70). A CsCl gradient established during spinning. After centrifugation the plasmid DNA, as well as the genomic DNA, RNAs and proteins, was visualized by UV-exposure of the tube. The plasmid DNA band was collected by suction through a needle with a 2 ml syringue. The ethidium bromide was extracted from the DNA by successive extractions with water-saturated butan-2-ol. Approximately 5 ml of DNA-containing phase was usually recovered and split between two tubes. To each 2.5 ml solution, 8 ml of 1X TE was added and the DNA precipitated by addition of 8 ml of propan-2-ol and incubation for 30 min at -20°C. After centrifugation the DNA pellet was resuspended in 500 µl of 1X TE and kept at -20°C for an indefinite time.

- Lysis buffer: 0.1% (v/v) Triton X-100, 63 mM EDTA pH 8.0, 50 mM Tris-Cl pH 8.0.

#### 2.4.3.2 Yeast genomic DNA preparation

Genomic DNA was routinely prepared from a 5 ml mid-logarithmic culture grown in YPDA medium (section 2.1.2.3). The culture was harvested by centrifugation for 5 min at 3000 rpm (Mistral 1000 centrifuge). The supernate was discarded, the cells resuspended in 0.5 ml of [1 M sorbitol, 0.1 M EDTA] solution containing 2 mg of lyticase. This cell suspension was transferred to a 1.5 ml Eppendorf tube and incubated

at 37°C for 1 h with continuous slow shaking until conversion to spheroplasts was complete. Spheroplasts were sedimented by centrifugation in a microfuge for 1 min. The supernate was discarded and the pellet gently resuspended in 0.5 ml of [50 mM Tris-HCl pH 7.4, 20 mM EDTA] solution. 50 µl of 10% (w/v) SDS was added and mixed well. The mixture was inbubated at 65°C for 30 min. 200 µl of 5 M potassium acetate was then added, the suspension mixed by inversion and placed on ice for 60 min. The suspension was then centrifuged in a microfuge for 5 min and the lysate transferred in a fresh Eppendorf tube. The DNA was finally precipitated in the presence of one volume of isopropanol. The mix was incubated at room temperature for 5 min and the DNA pelleted by briefly spinning in a microfuge. The supernate was pour off and the pellet air dried. The final DNA pellet was resuspended in 300 µl 1X TE buffer (section 2.1.3), 15 µl of a 1 mg/ml of RNase A solution was added and incubated at 37°C for 30 min. The DNA solution was then kept for an indefinite time at -20°C.

#### 2.4.3.3 Restriction digest of DNA

Restriction endonuclease digestion of DNA was performed in volumes of 10-20 µl. These contain the required quantity of DNA and the appropriate Boehringer Mannheim restriction buffer at 1X concentration. The restriction enzyme volume was kept below 10% of the digestion volume to avoid any star activity of the enzyme due to excess glycerol. The restriction digest was incubated at the temperature recommended by the supplier for a period of 2-10 h. The products of the digestion were either analysed directly by agarose gel electrophoresis, or extracted with phenol:chloroform then ethanol precipitated and finally resuspended in 1X TE buffer (section 2.1.3) in a suitable volume for further manipulations.

- Phenol:chloroform extraction mixture: these organic solvents were mixed with a volume ratio of 25:24 respectively with 1 volume of isoamyl alcohol.

### 2.4.3.4 Agarose gel electrophoresis

Agarose gel electrophoresis analysis of DNA was always performed using TAE buffer (section 2.1.3). Gels were prepared by melting the agarose (typically 0.7-1% (w/v)) in 1X TAE buffer. Ethidium bromide was added to a final concentration of 0.5 μg/ml for the visualisation of DNA. Samples to be analysed by gel electrophoresis were loaded directly using 6X Ficoll loading buffer.

- 6X Ficoll Loading Buffer: 0.25% (w/v) Bromophenol blue, 0.25% (w/v) Xylene cyanol blue, 15% (w/v) Ficoll.

### 2.4.3.5 Isolation of DNA from agarose gel slices

To isolate and purify DNA from agarose gels, the Qiagen gel extraction kit (Qiagen) was employed. QIAEX utilises a silica gel suspension which binds to DNA in high-salt solutions only. After several washes DNA was recovered by elution in low salt concentration buffer or water.

DNA fragments were separated on a TAE/agarose gel (section 2.4.3.2). The band of interest was detected by UV-transillumination (UVT 28M Herolab) on low power and cut out using a clean razor blade. The gel slice was transferred to an Eppendorf, the weight of the gel slice determined and three volumes of QX1 solution (Qiagen) added. 10 μl of QIAEX matrix was then added and the Eppendorf incubated at 50°C for 10 min until the gel slice had dissolved; the Eppendorf was vortexed 2-3 times during the incubation period in order to help the dissolution of the gel and the DNA-silicagel particles interaction. The QIAEX matrix was sedimented by centrifugation for 30 sec, the supernate removed and the pellet washed twice in 500 μl of QX1 solution. A last wash was done with 500 μl of PE solution (Qiagen), the pellet was air-dried for 10-15 min and the DNA eluted by resuspending the matrix in 20 μl 1X TE or water. Finally, the DNA-containing solution was separated from the QIAEX by centrifugation and the

supernate transferred to a new Eppendorf tube. The DNA was directly used for further manipulations or kept at -20°C.

### 2.4.3.6 Ligation of DNA molecules

Ligations were typically performed in a final volume of 10-20 µl. These contained 0.5-1.0 µg total DNA, 1X ligation buffer and T4 DNA ligase.

Cohesive DNA termini were ligated with 1 unit of ligase at RT for 1-2 h, the ligation of blunt-ended molecules was performed overnight at 4°C. Vector and insert DNA were typically present in a 1:2 or 1:3 ratio. 5 µl of the ligation reaction was used to transform competent *E. coli* cells (section 2.3.2).

- 1X ligation buffer: 300 mM Tris-Cl (pH adjusted to 7.5), 100 mM $MgCl_2$, 100 mM DTT, 10 mM ATP.

### 2.4.3.7 "Filling in" of recessed 3' termini

Klenow DNA polymerase I was used to fill-in recessed 3'termini generated by various restriction endonucleases or some DNA polymerases. Reactions were performed in a final volume of 20-50 µl containing 1X buffer (supplied by the manufacturer). All four deoxyribonucleotides were present at a concentration of 20 mM each with 2 units of klenow DNA polymerase I. The reaction was incubated for 1 h at 37°C and the unincorporated dNTPs removed by phenol/chloroform extraction (section 2.4.3.1). The DNA was ethanol precipitated (section 2.4.2) and resuspended in a suitable volume of 1X TE (section 2.1.3).

### 2.4.3.8 Amplification of DNA using the polymerase chain reaction

Usually 10-20 ng of circular plasmid DNA or up to 1 μg of genomic DNA were used as reaction templates. PCR was also done directly on yeast colonies simply by mixing the yeast cells with the reaction mix just prior starting the cycles. Transformant *E. coli* colonies were also screened by PCR: cells were boiled in 200-500 μl of sterile water for 5 min and 5-10 μl of the supernate was used as DNA template in a PCR reaction.

Specific DNA oligonucleotide primers were obtained commercially (Table 8) and used at a concentration of 0.1-1 μM. Deoxyribonucleotides were present at a concentration of 1 mM and $MgCl_2$ at 2.5 mM-4 mM. A 10X DNA polymerase reaction stock buffer was supplied by the manufacturer and diluted to 1X concentration in the reaction mix. *Pfu* or $Vent_R$ DNA polymerases were used when an accurate amplification of the DNA template was needed. Both enzymes possess a high-fidelity polymerase ability due to their 3'-5' proofreading exonuclease activity. *Taq* DNA polymerase was used when this high fidelity activity was not required.

PCR reactions were carried out in a Hybaid™ Thermal Reactor or a PTC-100™ Hot Lid (Genetic Research Instrumentation Ltd) apparatus. The three-steps cycle reaction was programmed according to the length of the desired product and the appropriate melting temperature of the primer/template duplex. A typical program was containing the following steps:

***Step I*:** Double-stranded DNA was denatured by heating, usually for 2 min at 92-94°C.
***Step II*:** Primers were annealed to their complementary sequences by briefly cooling down the temperature of the sample reaction. This annealing temperature depends on the length and the sequence of the primers. The following formula was used to estimate the annealing temperature of 20 bases long primers (or shorter) : $T = 2(A+T) + 4(G+C)$, where A, T, G and C were the numbers of those bases in the oligonucleotide (Kidd *et al.* 1995).

***Step III***: Polymerisation occurred, the annealed primers were extended by the DNA polymerase. Usually a temperature between 70 and 75°C was used for this step. one minute polymerisation was utilised per Kb to be amplified.

2.4.3.9 Labelling DNA fragments by random priming

DNA fragments of cloned genes were radiolabelled by the random priming method of Feinberg and Vogelstein (1983, 1984).

The DNA to be labelled was gel-purified as described in section 2.4.3.5. 50 ng to 100 ng of this DNA was made up to a 35 µl volume with water, denatured by boiling for 5 min, then cooled down on ice. The random priming reaction was prepared as follows and incubated overnight at room-temperature:

| | |
|---|---|
| **Oligo Labelling Buffer (OLB)** | 10 µl |
| **BSA (20 mg/ml; Boehringer Mannheim)** | 2 µl |
| **[α-$^{32}$P] dCTP (10 mCi/ml)** | 3 µl |
| **Denatured DNA in water** | 35 µl |
| **Klenow DNA Polymerase I** | 2 units |

- *OLB*: consists of a 1:2.5:1.5 mixture of solutions A:B:C prepared as follows:

*Solution A*: 18 µl β-mercaptoethanol, 5 µl of each 0.1 M dNTPs prepared in 1 ml solution O. Stored at -20°C.

*Solution B*: 2 M Hepes buffer (pH adjusted to 6.6 with KOH). Stored at 4°C.

*Solution C*: Random hexadeoxyribonucleotides (Pharmacia) at 90 OD Units/ml in sterile water. Stored at -20°C.

*Solution O*: 1.25 M Tris-HCl pH 8.0, 0.125 M $MgCl_2$. Stored at 4°C.

Random primed DNA was separated from unincorporated nucleotides using a NAP-5 gel filtration column (Pharmacia) according to the manufacturer's instructions.

### 2.4.3.10 End-labelling of oligonucleotides

Oligonucleotides were labelled in a 20 µl reaction at 37°C for 45 min using New England Biolab T4 Polynucleotide Kinase as follows:

| | |
|---|---|
| **Oligonucleotide** | 10 pmoles |
| **10X T4 PNK buffer (supplied by NEB)** | 2 µl |
| **[$\gamma$-$^{32}$P] ATP (5 000 Ci/mmole, 10 mCi/ml)** | 2 µl |
| **T4 PNK** | 2 units |

The total volume was adjusted to 20 µl with water.

Radiolabelled oligonucleotides were separated from unincorporated nucleotides using the gel filtration NAP-5 column (Pharmacia) according to the supplier's instructions.

### 2.4.3.11 Southern blotting

DNA was digested with the desired restriction enzymes and electrophoresed through an agarose gel (sections 2.4.3.4). The gel was first immersed in denaturing buffer with gentle agitation for 30 min, then transferred to neutralisation buffer for a further 30 min.

#### *Transfer of DNA from agarose gels to nylon membranes*

Transfer of DNA fragments was as described by Southern (1975). Hybond™-N nylon membrane (Amersham) and four sheets of Whatman 3MM paper were cut to the same size as the gel. One sheet of blotting paper was saturated with 20X SSC (section 2.1.3) and placed over a clean glass plate, arranged so that it hung over the edge of the

plate with the ends of the paper in a plastic tray containing 20X SSC. The pre-treated gel was placed on top of the saturated Whatman 3MM paper and Saran wrap carefully placed on top of regions of the paper left uncovered by the gel to prevent unnecessary evaporation. The nylon membrane was placed on top of the gel, taking care to ensure no air bubbles formed between the gel and the membrane. The four sheets of pre-cut Whatman 3MM paper were saturated in 20X SSC and placed on top of the membrane, again taking care to prevent the formation of bubbles between the layers. On top of this, 2–4 cm of dry paper towels were arranged, and the whole structure weighted to provide even pressure. Transfer was allowed to take place overnight. Once the transfer was complete, the membrane was briefly rinsed in 20X SSC, blotted dry, and UV-irradiated in a Stratagene UV Stratalinker™ using the "autocrosslink" setting (1200 μjoules, 254 nm) to immobilise the DNA to the filter.

### *Hybridisation of randomly-labelled probes to nylon membranes*

Hybridisations were performed in Hybaid 'Hybridiser ovens' using an adaptation of the method of Church and Gilbert (1984).

The nylon membrane was pre-hybridised in 20 ml of hybridization buffer for at least 3 h at 60°C-65°C to prevent non-specific hybridisation of the probe to the membrane. Fresh hybridization buffer was added to the membrane immediately before the addition of the radiolabelled probe. The labelled probe (section 2.4.3.9) was denatured by heating to 100°C for 5 min, then added to the hybridization buffer and the incubation continued overnight at 60°C (or lower temperature depending on the annealing temperature of the target DNA-prode hybrid); this hybridisation temperature was increased or decreased depending on the desired stringency conditions. The following morning the probe was decanted off and stored at –20°C for possible re-use. The membrane was then washed with washing buffer for 20 min at 60°C; this temperature was varied to suit the stringency required. This was repeated four times. The membrane was then blotted dry, placed between Saran wrap, and the result of the experiment visualised by autoradiography.

- Denaturing Buffer: 0.5 M NaOH, 1.5 M NaCl.

- Neutralisation Buffer: 1.5 M NaCl, 0.5 M Tris-HCl pH 7.5, 1 mM EDTA.

- Hybridization buffer: 7% (w/v) SDS, 1 mM EDTA pH 8.0, 0.5 M Sodium phosphate buffer pH 7.2.

- Washing buffer: 5% (w/v) SDS, 1 mM EDTA pH 8.0, 40 mM Sodium phosphate buffer pH 7.2.

#### 2.4.3.12 Site directed mutagenesis

The following mutagenesis technique was based on a method described by Kunkel (1987): The DNA sequence to be mutated was cloned into the phagemid vector pBluescript $KS^{+/-}$ (Table 9). The construct was transformed into the *E. coli* strain CJ236 (Table 6). This strain was a double mutant *dut*$^-$ *ung*$^-$ bacterium and produces nascent DNA carrying a number of uracils in thymidine positions. Single-stranded copies of DNA were isolated from this transformed strain. A mutagenic oligonucleotide complementary to the region to be mutated was designed (Table 8) and hybridized (except for the altered region) to the single-strand copy of the plasmid DNA. A complementary strand was then synthesized by T4 DNA polymerase using the oligonucleotide as a primer. T4 DNA ligase was then used to seal the new strand to the 5' end of the oligonucleotide. The double stranded DNA, containing a mismatch at the mutated sequence was transformed into *E. coli* strain DH5αF' (Table 6), resulting in two classes of progeny: The parental one (containing the wild-type DNA sequence and a number of uracils) and the mutated one (containing the mutated sequence and no uracils). Due to the wild-type uracil N-glycosylase activity of DH5αF', the parental uracil-containing strand was inactivated with high efficiency leaving the mutated strand to replicate and an enrichment for the cells carrying the mutated DNA.

### *Propagation of uracil-containing phagemids*

Competent CJ236 cells (Table 2.3) were transformed with phagemid DNA (pBluescript $KS^{+/-}$ containing the insert to be mutated; Table 9) as described (section 2.3.2). An isolated colony of CJ236 containing the phagemid was used to inoculate 20 ml of LB-Cm/Amp medium (section 2.1.2.1, Table 4) and the culture grown overnight with shaking at 37°C. The following morning, 50 ml of YT-Amp medium (section 2.1.2.1) was inoculated with 1 ml of the overnight culture and the cells grown at 37°C with shaking. When $OD_{600}$= 0.3 (this corresponds to approximately $1x10^7$ cfu/ml, the helper phage M13K07 was added in the culture to obtain a M.O.I. of 20 phage/cell. The culture was incubated for an additional 1 h, then 70 µl of the 50 mg/ml kanamycin stock (Table 4) were added and the incubation continued for a further 4-6 h.

### *Harvesting phagemid*

Cultures (30 ml) were transferred to a sterile 30 ml corex tube and the cells pelleted by centrifugation at 12000 rpm for 15 min (4°C; Beckman centrifuge JA 25-50 rotor). The phagemid-containing supernate was transferred to a fresh corex tube and 150 µg RNase A were added, the contents were mixed by inversion and incubated at room temperature for 30 min. A quarter volume of 3.5 M ammonium acetate/20% (w/v) PEG-6000 was added, the contents of the tube were mixed thoroughly and kept on ice for 30 min. Phagemids were collected by centrifugation at 12000 rpm for 15 min, the supernate discarded, the pellet drained well and resuspended in 200 µl of high salt buffer (300 mM NaCl, 100 mM Tris-HCl pH 8.0, 1 mM EDTA). The prepared phagemid suspension was transferred to an Eppendorf tube and chilled on ice for 30 min. The suspension was centrifuged for 2 min and the phagemid-containing supernate transferred to a fresh Eppendorf tube. The phagemid DNA was either extracted immediately or stored as phagemid particle for up to one week at 4°C.

### *Titering phagemid stock*

Since the phagemids contain uracil in their DNA, they should survive far more readily in a bacterium missing an active uracil N-glycosylase (*ung*$^-$; CJ236) than in one with an active enzyme such as DH5αF'.

Cultures of CJ236 and DH5αF' were grown overnight at 37°C in LB-medium (section 2.1.2.1). The following morning, 50 ml of fresh LB-medium was inoculated with 1 ml of CJ236 culture or 0.5 ml DH5αF' culture. These were incubated with shaking at 37°C until $OD_{600}$ 0.3 to 0.35; if one of the cultures reached this OD value before the other it was kept on ice. Once the desired OD was reached, 10 µl of uracil-containing phagemid stock was added and each culture incubated for a further 2 h at 37°C. An aliquot of DH5αF' culture was diluted 10 and 100 fold, then 50 µl of each dilution and undiluted culture spread onto LB-Amp plates. An aliquot of CJ236 culture was diluted $10^3$-, $10^4$-, and $10^5$-fold and 50 µl of each of those dilutions spread onto LB-Amp plates. The plates were incubated overnight at 37°C.

***Extraction of the phagemid DNA***

DNA was recovered from the phagemid stock by successive extractions with phenol, phenol:chloroform and chloroform, followed by precipitation with ethanol in the presence of 0.1 volume of 7.8 M ammonium acetate. The DNA pellet was washed carefully with 70% (v/v) ethanol, dried and resuspended in 20 µl 1X TE.

***Synthesis of the mutagenic strand***

Synthesis of the mutagenic strand was achieved by priming synthesis from the single-stranded uracil-containing DNA template with the oligonucleotide containing the sequence of the desired mutation (Table 8). This synthesis can be divided into three steps.

*Phosphorylation of mutagenic oligonucleotide*. Up to 200 pmoles of mutagenic oligonucleotide was phosphorylated by preparing the following mix:

| | |
|---|---|
| **200 pmole Mutagenic oligonucleotide** | variable volume |
| **1 M Tris-Cl, pH 8.0** | 3 µl |
| **0.2 M $MgCl_2$** | 1.5 µl |
| **0.1 M DTT** | 1.5 µl |
| **1 mM ATP** | 13 µl |

| | |
|---|---|
| **Sterile water** | Adjusted to a final volume of 30 µl |

To this mix 4.5 units of T4 Polynucleotide Kinase was added and the mixture incubated at 37°C for 45 min. The reaction was stopped by heating at 65°C for 10 min. The phosphorylated oligo was diluted to 6 pmole/µl with TE (section 2.1.3) and stored frozen.

*Annealing of mutagenic oligonucleotide to template DNA*. The following reaction was prepared in an Eppendorf:

| | |
|---|---|
| **10 pmole Mutagenic oligonucleotide** | variable volume |
| **0.2 µg Uracil-containing template DNA** | variable volume |
| **10X Annealing buffer** | 1 µl |
| **Sterile distilled water** | Adjusted to a final volume of 10 µl |

- 10X Annealing buffer: 200 mM Tris-Cl pH 7.5, 20 mM $MgCl_2$, 500 mM NaCl.

The mixture was incubated for 5 min at 70°C. The reaction was allowed to cool slowly to room temperature over 40-60 min, then stored on ice. A control reaction containing all of the above ingredients but lacking the mutagenic primer was also set up to test for non-specific endogenous priming, caused by contaminating nucleic acids in the template preparation.

*Synthesis of the complementary strand*. To the mixture prepared previously, 1 µl of 10X synthesis buffer, 3 units of T4 DNA ligase and 1 unit of T4 DNA polymerase were added. The reaction was incubated for 5 min on ice to stabilize the primer by initiation of DNA synthesis under conditions that favour binding of the primer to the template. The mixture was then incubated at 25°C for 5 min and finally at 37°C for 90 min. After this time, 90 µl of stop buffer were added and the reaction stopped by freezing. This mixture can be kept at -20°C for at least one month. The successful production of circular double standed DNA can be analysed by agarose gel electrophoresis (section

2.4.3.4). A 9 µl aliquot of each *in vitro* mutagenesis reaction (with and without primer added) were loaded on an agarose gel: The conversion of the single stranded DNA to a slower migrating double stranded form should be noticed only when the primer is added in the reaction mix.

- 10X Synthesis buffer: 4 mM each dNTPs, 7.5 mM ATP, 175 mM Tris-Cl pH 7.4, 37.5 mM $MgCl_2$, 5 mM DTT.

Stop buffer: 10 mM Tris-Cl pH 8.0, 10 mM EDTA.

***Transformation of the reaction mix into E. coli* DH5αF' *and screening for the mutated clones***

*In vitro* synthesised mutagenic double stranded DNA (5 µl) was used to transform 150 µl of competent DH5αF' cells (section 2.3.2). As a control another transformation was done using 3 µl of the no primer reaction. Plates were incubated at 37°C overnight. Double-stranded DNA sequence analysis was used to identify the desired mutant clones (section 2. 4. 3. 13).

2.4.3.13 DNA sequencing

All sequencing reactions were carried out using the Sequenase Version 2.0 sequencing kit (USB) according to the manufacturers' instructions. Templates for the reactions were either double-stranded DNA plasmid clones or single-stranded preparations of phagemid.

***Preparation of single-stranded plasmid DNA templates***

DH5αF' cells carrying the DNA of interest to be sequenced on a phagemid were grown at 37°C overnight to stationary phase in selective LB-medium (section 2.1.2.1). A phage-infected culture was prepared by adding 50 µl of the overnight culture to 2 ml of LB-medium in a sterile 20 ml bijoux tube, growing the cells for 30 min at 37°C and

then adding one tenth of a suspension of M13KO7 helper phage ($1x10^{11}$ pfu/ml). The infected culture was then incubated for 5-6 h at 37°C with constant vigorous agitation. 1.5 ml of the infected culture were transferred to an Eppendorf and centrifuged for 5 min at 4°C. 1 ml of the supernate was then transferred to a fresh Eppendorf and 200 µl of 20% PEG 8000 in 2.5 M NaCl [prepared fresh from sterile 40% (w/v) PEG 8000 and 5 M NaCl solutions] was added, the solutions mixed by vortexing and incubated at room temperature for 15 min. Bacteriophage particles were sedimented by centrifugation for 10 min at 4°C and the supernate discarded. The sample was briefly re-centrifuged, any residual supernate carefully removed and the pellet resuspended in 100 µl of 1X TE (pH 8.0). Single-stranded DNA was recovered from the phage-suspension by successive extractions with phenol, phenol:chloroform and chloroform, and finally ethanol precipitated. The DNA pellet was washed in 70% (v/v) ethanol, dried and dissolved in 20 µl of sterile water. 1 µg of single stranded DNA was used per sequencing reaction.

***Preparation of double-stranded plasmid DNA templates***

Plasmid DNA (3-5 µg; section 2.4.3.1) to be sequenced was dissolved in distilled water such that the total volume was 20 µl. The DNA was denatured by adding 0.1 volume of [2 M NaOH, 2 mM EDTA] and incubated 30 min at 37°C. The denaturation reaction was neutralised by the addition of 0.1 volume of 3 M sodium acetate (pH 5.2), and the DNA precipitated with 2-4 volumes of 100% ethanol. The denatured DNA-template was recovered by centrifugation for 10 min at 4°C, the pellet washed with 70% (v/v) ethanol and dried. The DNA pellet was resuspended in 7 µl of water and ready to be sequenced.

Three steps (annealing template and primer, labelling reaction and termination reaction) were followed for DNA sequencing:

***Primer annealing reaction***

Single- or double-stranded DNA (7 µl) were mixed with 2 µl of 5X Sequenase buffer (Sequenase Version 2.0 sequencing kit) and 1 µl of sequencing primer (1 pmole). This annealing mix was heated at 65°C for 2 min then cooled down slowly to room temperature over 30-40 min. The mixture was then centrifuged briefly and left on ice until used.

***Labelling reaction***

A labelling reaction mixture was set up in a separate Eppendorf as follows, adding the individual components in the following order:

| | |
|---|---|
| **dGTP labelling mix** (Sequenase Kit; diluted as required) | 2 µl |
| **0.1 M DTT** | 1 µl |
| **[$^{35}$S] dATP (10 µCi/µl)** | 0.5-1 µl |
| **Sequenase Enzyme** (diluted) | 2 µl |

5-6 µl of this labelling reaction mix were added to the 10 µl annealed template/primer reactions stored on ice, the solutions gently mixed then incubated at room temperature for 2-5 min.

The sequenase DNA polymerase was diluted eight-fold in glycerol enzyme dilution buffer (USB) just prior use.

For standard sequencing reactions the dGTP labelling mix was 5-fold diluted in water, for sequencing reactions close to the primer this mix was 15-fold diluted. To read sequences close to the primer, Mn buffer (USB) was added in the labelling cocktail (1 µl of 1 M stock solution) to emphasize bands close to the primer.

***Termination reaction***

For each template to be sequenced, four Eppendorf tubes were labelled 'A', 'C', 'G' and 'T' respectively and 2.5 µl of each of the four ddNTP termination mixes were placed into the appropriately labelled Eppendorfs; these tubes were pre-incubated at

37°C until use. Termination reactions were initiated by adding 4.5 µl of the labelling reaction to each of the pre-warmed ddNTP mixes. The solutions were gently mixed, the Eppendorfs placed back at 37°C for 5 min, then 4 µl of Sequenase stop-solution (USB) was added to each Eppendorf. Reactions were stored at –20°C until required for electrophoresis.

The sequencing reactions were heated at 85°C for 3 min immediately before loading onto a 6% (w/v) denaturing polyacrylamide gel (section 2.4.4.2). Gels were run at 33 watts for 4-10 h then dried on a vacuum gel drier (Hybaid) at 80°C for 2 h. Sequence was visualised by autoradiography (section 2.2.2).

### 2.4.4 Ribonucleic acid methods

#### 2.4.4.1 *In vitro* RNA transcription

T7 and SP6 polymerase-directed *in vitro* transcription reactions were performed essentially as described by Yisraeli *et al.* (1989) and Milligan *et al.* (1989). Three different types of DNA templates were used to transcribe RNA molecules *in vitro*:

- Plasmid DNAs (prepared according to the large scale cesium chloride preparation method described in section 2.4.3.1). For run-off transcription, plasmids were linearized by endonuclease digestion (section 2.4.3.2), phenol/chloroform extracted and ethanol precipitated. The DNA template was finally resuspended in water in a suitable volume to give a DNA concentration of 1 mg/ml.

- PCR DNA templates (section 2.4.3.8): Primer oligos described in Table 8 were used in the PCR reaction.

- DNA oligonucleotides containing at their 3' end the T7 promoter sequence: This oligo was annealed to the T7 primer (Table 8). The T7 transcription was initiated from this [primer-DNA oligo] hybrid.

### *Production of unlabelled RNA molecules*

Usually a 100 µl reaction was performed in order to produce a large quantity of RNA. This reaction was prepared as follows and incubated for 2 h at 37°C:

| | |
|---|---|
| **10X Transcription buffer** | 10 µl |
| **10X rNTPs mix** | 10 µl |
| **DNA template** | 10 µg |
| **RNasin** | 3 µl |
| **T7/SP6 RNA Polymerase** | 350 units |
| **Sterile water** | To a final volume of 100 µl |

When the transcription template was a DNA oligo, between 5 and 10 pmole of this oligo was mixed with 50-100 pmole of T7b primer (Table 8); other transcription components were added as indicated above.

- 10X T7 RNA polymerase-mediated transcription buffer: 400 mM Tris-Cl pH 7.5, 20 mM spermidine, 100 mM DTT, 60 mM $MgCl_2$, 100 mM NaCl.

- 10X SP6 RNA polymerase-mediated transcription buffer: 400 mM Tris-Cl pH 7.5, 40 mM spermidine, 100 mM DTT, 100 mM $MgCl_2$, 100 mM NaCl.

- 10X rNTPs mix: 10 mM each ATP, CTP, UTP and GTP in sterile water.

The reaction was stopped by addition of 2 µl of 0.5 M EDTA. Phenol/chloroform (100 µl) was added to the reaction mix, after vortexing and centrifugation in a microfuge, the aqueous phase containing the RNA was extracted and transferred to a fresh Ependorf tube. The RNA was precipitated with ethanol, after centrifugation, the RNA pellet was resuspended in 20 µl of sterile water and 20 µl of 100% formamide loading dye was added. The sample was boiled for 5 min and loaded on an Acrylamide/Urea/TBE gel (section 2.4.4.2). The RNA molecule was detected by UV-shadowing and isolated from the abortive initiation products according to the method described in section 2.4.4.3.

### *Production of $^{32}$P-radiolabelled RNA molecules*

The commonly $^{32}$P-radiolabelled RNA molecules were the substrate pre-mRNA and the U6 snRNA. Those RNA molecules were produced with a different specific activity because of their different use.

*Transcription of $^{32}$P-radiolabelled pre-mRNA*: When the substrate was prepared for subsequent *in vitro* splicing reaction (section 2.6.2), a 10 µl reaction was set-up in a sterile Eppendorf tube and incubated for 30 min at 37°C. The following components were present:

| | |
|---|---|
| **DNA Template** | 1 µg |
| **10X T7or SP6-Transcription buffer** | 1 µl |
| **RNasin** | 0.3 µl |
| **10 mM each ATP, CTP and GTP in sterile water** | 1 µl |
| **250 µM cold UTP** | 0.5 µl |
| **T7 or SP6 RNA polymerase** | 35 units |
| **$^{32}$P-αUTP (20 µCi/µl, 800 Ci/mmole)** | 2 µl |
| **Sterile water** | To a final volume of 10 µl |

Typically the DNA template used was the actin gene cloned in the plasmid p283 (Table 2.6), linearized with the *Bam*HI restriction enzyme. The RP51 pre-mRNA was also transcribed and was linearized either with *Bam*HI to produce a pre-mRNA with a 100 nucleotides long poly(A) tail, or linearized with *Eco*RI endonuclease to produce a pre-mRNA lacking this poly(A) tail.

RNA molecules produced according to the recipe described above possessed a specific activity of 230 Ci/mmole.

Susbtrates prepared for UV-crosslinking experiments (section 2.7) required a higher specific activity, therefore only 20 pmole of cold UTP were added instead of 125 pmole recommended in the given protocol. This lead to the transcription of a pre-mRNA with a specific activity of 570 Ci/mmole.

Following transcription, an equal volume of 100% formamide loading dye was added (section 2.4.4.3). The sample was boiled for 5 min and then loaded on an Acrylamide/Urea/TBE gel. The radiolabelled U6 snRNA was detected by autoradiography (section 2.2.2) and subsequently gel purified according to the method described in section 2.4.4.3.

*Transcription of $^{32}$P-radiolabelled-U6 snRNA*: This snRNA was subsequently used for reconstitution and UV-crosslinking reactions (sections 2.6.2.3 and 2.7), therefore a high specific activity-containing RNA was needed. The following reaction mixture was prepared in a sterile Eppendorf tube and incubated for 1 h at 37°C:

| | |
|---|---|
| **U6 DNA Template** | 1 µg |
| **10X T7-Transcription buffer** | 1 µl |
| **RNasin** | 0.3 µl |
| **10 mM each UTP, CTP and GTP in sterile water** | 1 µl |
| **100 µM ATP** | 0.2 µl |
| **T7 RNA polymerase** | 35 units |
| **$^{32}$P-αATP (10 µCi/µl, 400 Ci/mmole)** | 4 µl |
| **Sterile water** | To a final volume of 10 µl |

The wild-type U6 snRNA gene template was provided by the plasmid pUC12U6 (Table 9), linearized with the *Dra*I restriction enzyme. Mutated U6 snRNAs were also produced using template DNA plasmids pU6C48A, U6A51U, U6A59U and U6G60C (Table 9) *Dra*I digested. Shorter U6 snRNA molecules (lacking 17 nucleotides at the 3' end of the wild-type sequence) were transcribed using as DNA template the pUC12U6 plasmid digested with *Fok*I restriction enzyme. The U6-ribozyme was transcribed from pU6ribo linearized with *Eco*RI.

Although the transcription recipe given was for a 10 µl reaction, it was usually scaled-up in order to produce the larger quantities of U6 snRNA necessary for several U6 snRNP reconstitution reactions (section 2.6.2.3). The transcribed U6 snRNA possesses a specific activity of 330 Ci/mmole.

Following transcription, an equal volume of 100% formamide loading dye was added (section 2.4.4.3). The sample was boiled for 5 min and then loaded on a 8% Acrylamide/Urea/TBE gel. The radiolabelled U6 snRNA was detected by autoradiography (section 2.2.2) and subsequently gel purified according to the method described in section 2.4.4.3.

### 2.4.4.2 Acrylamide gel electrophoresis of RNA

Six or eight percent acrylamide/8M urea/1X TBE gels were prepared by mixing ready to use Ultrapure sequagel solutions, "Complete Buffer Reagent" and APS in a volume recommended by the manufacturer. For other percentages of acrylamide gels, a 40% (w/v) acrylamide/2.1% (w/v) bisacrylamide stock solution was diluted according to the final percentage required and mixed with powder urea and 10X TBE to finally obtain an acrylamide/8M urea/1X TBE solution. APS and TEMED were then added to a final concentration of 1.25 % (w/v) and 0.08% (w/v) respectively.

- Acrylamide gel electrophoresis was performed using a 1X TBE buffer (section 2.1.3).

- Before loading the acrylamide/urea/TBE gel, RNA samples were boiled for 5 min in 100% formamide loading dye.

◆ 100% Formamide loading dye: 40 mM EDTA, 0.3% (w/v) Xylene cyanol, 0.3% (w/v) Bromophenol blue, 93% (v/v) deionised formamide.

### 2.4.4.3 SnRNA blotting

Co-immunoprecipitated snRNAs were prepared as indicated in section 2.9 and loaded on a 6% Urea/TBE/Acrylamide gel (section 2.4.4.2) using the glass plates 'Base-Runner System', assembled as instructed by the manufacturer (IBI, Kodak). As a control, a total yeast RNA sample was prepared by treating 5 µl of yeast extract with

proteinase K (sections 2.9 and 2.7.1). The total RNAs were extracted with phenol/chloroform and precipitated with ethanol. The RNA pellet was resuspended in 5 µl of water and an equal volume of 100% formamide loading dye was added (section 2.4.4.2). After boiling the sample was loaded along with the co-precipitated snRNAs on the denaturing polyacrylamide gel. The gel was run at 33 watts until the bromophenol blue dye was 2 cm from the bottom of a plate of 43 cm height. The gel system was disassembled and the glass plates separated. The top seven centimetres of the gel, containing the U1 and U2 snRNAs, was cut and the RNA electrophoretically transferred on to a Hybond™-N nylon membrane (Amersham). From the remaining gel, the next 16 cm were cut and transferred to a nylon membrane: This contains the U4, U5 and U6 snRNAs. The RNA transfer was achieved by using the same procedure as indicated in section 2.5.5, except that the transfer buffer used was a 0.5X TBE solution and the transfer was performed at 60 volts for 1 h.

Once the transfer was complete, the nucleic acids were immobilised on the membrane by UV-irradiation of the filter in a Stratagene UV Stratalinker™ using the "autocrosslink" setting (1200 µjoules, 254 nm).

***SnRNA hybridisation***

Hybridisations were performed in Hybaid 'Hybridiser ovens' using an adaptation of the method of Church and Gilbert (1984): The nylon membrane was pre-hybridised in 20 ml of SES1 Buffer for at least 3 h at 65°C to prevent non-specific hybridisation of the probe to the membrane. Fresh SES1 buffer was added to the membrane and also the radiolabelled probe. The labelled probes consist of DNA oligos complementary to U1, U2, U4, U5 and U6 snRNAs (Table 8), $^{32}$P-end labelled according to the method indicated in section 2.4.3.10. The hybridisation temperature was adjusted to 5°C below the melting temperature of the oligonucleotide/snRNA hybrid. The hybridisation was carried for at least 6 h, then the hybridisation solution was poured off and stored at -20°C for possible re-use. Filters were washed three times for 30 min with fresh 20 ml of SES2 buffer each time. Usually the washing temperature was the same as the hybridization one. The radioactive activity of the filter was checked by monitoring it with a Geiger counter, and extra washes were carried out if too much radioactive

background was detected. The radioactive background was kept as low as possible. Finally, the membrane was placed between Saran wrap avoiding any air bubbles, and the result of the experiment visualised by autoradiography (section 2.2.2).

- SES1 buffer: 7% (w/v) SDS, 1 mM EDTA pH 8.0, 0.5 M Sodium phosphate buffer pH 7.2.

- SES2.buffer: 5% (w/v) SDS, 1mM EDTA pH 8.0, 0.5 M Sodium phosphate buffer pH7.2.

***Membrane stripping for reprobing***

For successful removal of probes, membranes were never allowed to dry during or after hybridisation and washing. A boiling solution of 0.1% (w/v) SDS was poured onto the Hybond™-N nylon membrane to be stripped and allowed to cool to room temperature. The membrane was blotted dry and stored between Saran wrap for future use.

2.4.4.4 Isolation and purification of RNA molecules from acrylamide gel

RNA molecules were resolved on an Acrylamide/Urea/TBE denaturing gel (section 2.4.4.1) and visualised either by UV-shadowing or detected by autoradiography when $^{32}$P-radiolabelled. The located band was cut out with a sterile razor blade, and depending on the RNA molecule size to be isolated two different methods of purification were followed:

***For short RNA molecules (below 35 bp)***

The RNA was passively eluted from the gel slice by incubating overnight at 4°C in 400 µl elution buffer. The following day the elution buffer containing the eluted RNA was transferred in a fresh sterile Eppendorf tube. The RNA was phenol:chloroform extracted and ethanol precipitated (section 2.4.3). The RNA pellet was resuspended in a suitable volume of water. The concentration was estimated by

spectrophotometry (section 2.4.1) or by Cerenkov counting (section 2.2.3) for a $^{32}$P-radiolabelled molecule.

- Passive elution buffer: 0.5 M NaOAc (pH 5.2), 1 mM EDTA, 2.5% (v/v) phenol:chloroform.

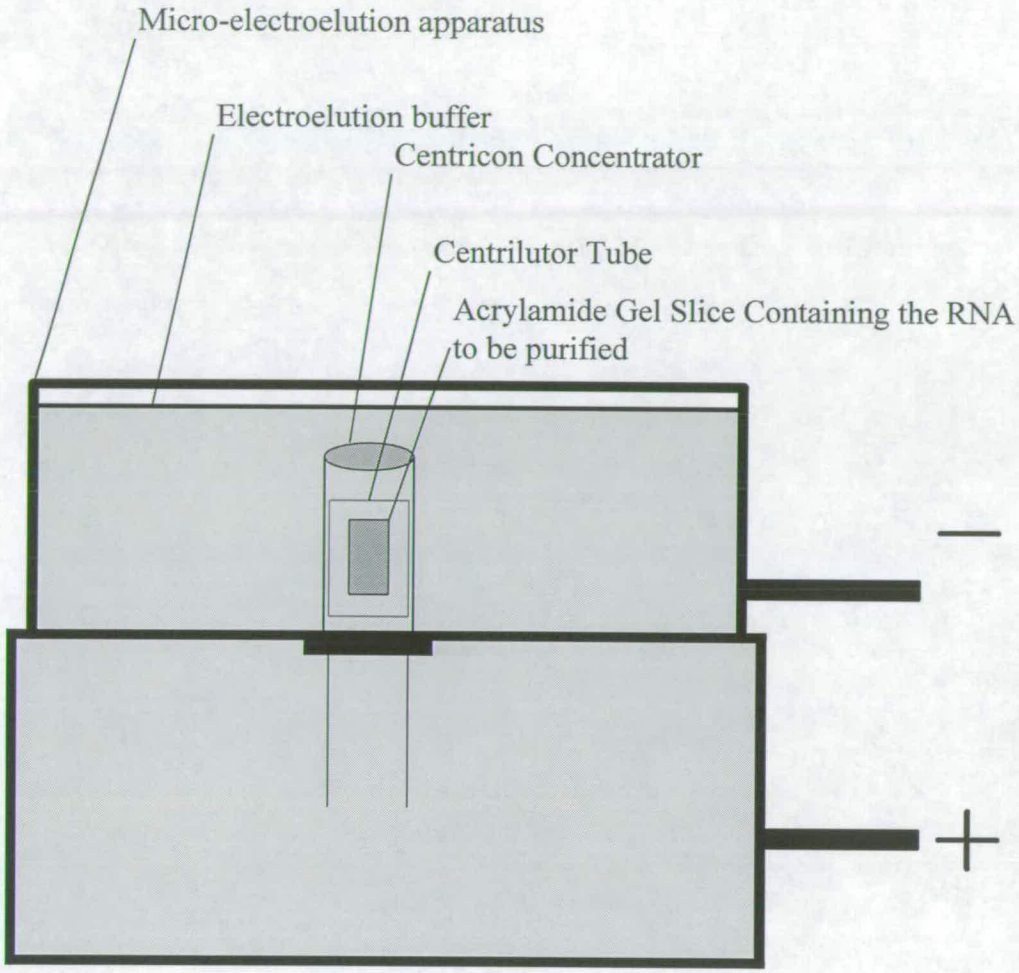
***For long RNA molecules (above 35 bp)***

The gel slice was transferred to a Centrilutor tube (Amicon) and the tube inserted in the sample reservoir of the Centricon concentrator C-10, C-30 or C-100 (Amicon) depending of the length of the RNA to be purified. The Centricon concentrator and the Centrilutor tube were assembled in a Micro-electroeluter (Amicon) apparatus and the electroelution chambers were filled with electroelution buffer. The RNA was electroeluted at 200 Volts for 1 h, then the electroelution system was disassembled and the Centrilutor tube removed from the Centricon concentrator. The electroelution buffer retained in the sample concentrator reservoir was concentrated by centrifugation according to the manufacturers'instructions. Finally, to the concentrated retentate, an equal volume of phenol:chloroform was added. After vortexing and centrifugation in a microfuge, the aqueous phase containing the RNA was extracted and transferred to a fresh Ependorf tube. The RNA was precipitated with ethanol, after centrifugation, the RNA pellet was resuspended in a suitable volume of water. The RNA concentration was estimated either by spectrophotometry (section 2.4.1) or by Cerenkov counting (section 2.2.3).

- Centricon concentrators C-10, C-30 and C-100 differ by the molecular weight cut-off of the polycarbonate membrane used to retain the molecules. This cut-off was respectively 30, 60 and 125 nucleotides long RNA molecules.

- Electroelution buffer: 1X TBE, 0.1% (w/v) SDS.

**RNA Electroelution Diagram.**

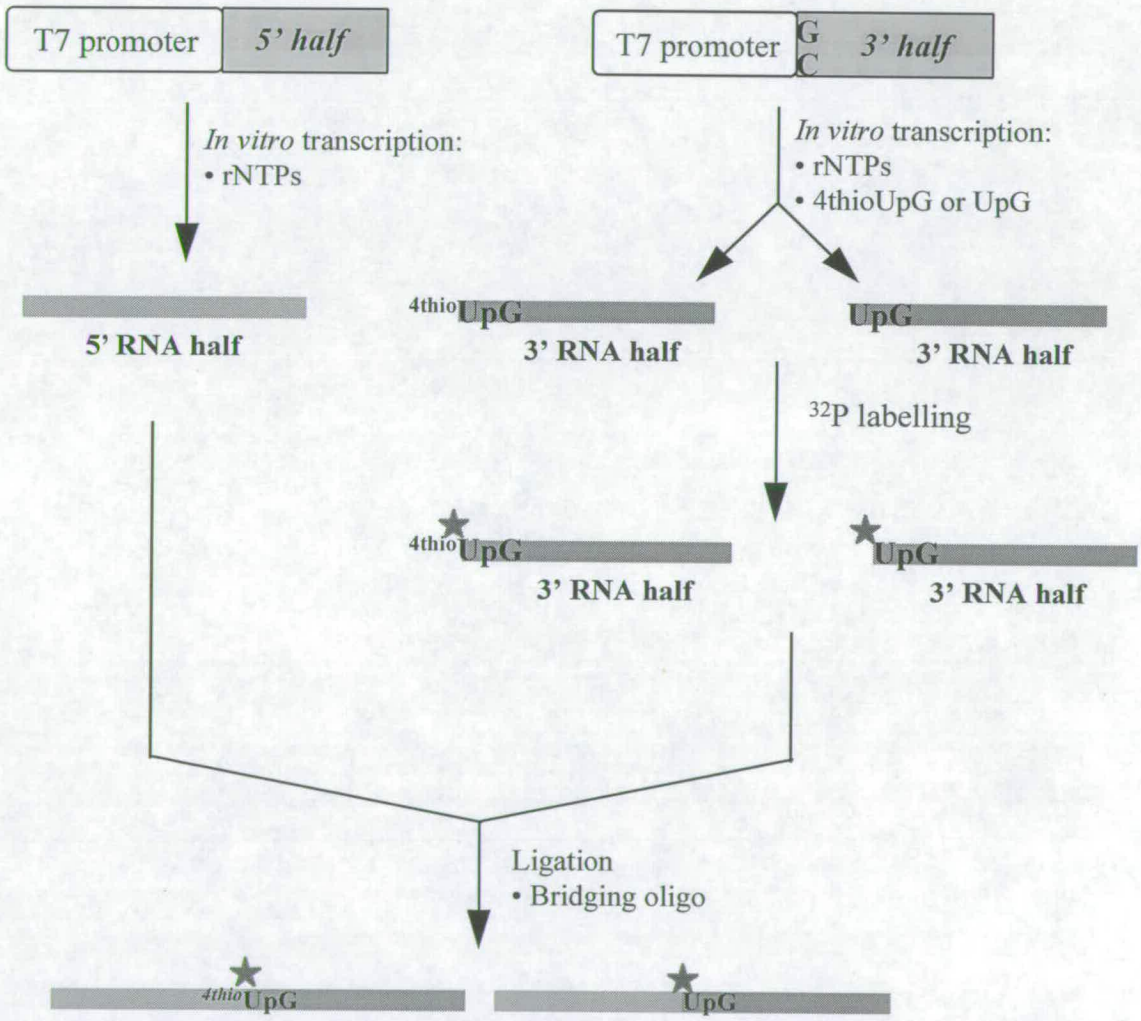



2.4.4.5 Site specific incorporation in RNA molecule of photoactivatable $^{32}$P-UpG or $^{32}$P-4thioUpG

The method used was described by Sontheimer (1994) and Newman *et al* (1995). Three steps were involved for the preparation of such RNA molecules:

### *Production of the RNA moieties*

The RNA molecule to be radiolabelled was constructed from two halves split at the site of the 4-thioU incorporation. This site lies obligatorily within a UG dinucleotide. The 5' and 3' halves were transcribed from a PCR product containing at its 5' end the T7 RNA polymerase promoter region or, for the synthesis of short RNA molecules (< 40 nucleotides), the transcription was initiated from a T7 primer annealed to a DNA oligo template. 4-thioUpG or UpG were used to prime the transcription of the 3' half.




Site specifically radiolabelled RNAs

Two transciption recipes were followed whether the 5' or the 3' half were produced:

*Production of the 5' half*: The transcription was performed in a 250 µl reaction and incubated for 2 h at 37°C. The reaction was prepared as follows:

| | |
|---|---|
| **10X Transcription buffer** | 25 µl |
| **100 mM $MgCl_2$** | 48 µl |
| **30% (w/v) PEG 8000** | 67.5 µl |
| **DNA template** | 7 µg |
| **10 mM each rNTPs** | 50 µl |
| **RNasin** | 5 µl |
| **T7 RNA polymerase** | 420 units |
| **Sterile water** | to a final volume of 250 µl |

- 10X Transcription buffer: 400 mM Tris-Cl pH 7.5, 10 mM spermidine, 50 mM DTT, 0.1% (v/v) Triton X-100.

The transcription reaction was stopped by adding 25 µl of 0.5 M EDTA. Phenol/chloroform (250µl) was added to the reaction mix, after vortexing and spin in a microfuge, the aqueous phase containing the RNA was extracted and transferred to a fresh Ependorf tube. RNA was precipitated with ethanol. The RNA pellet was resuspended in 15 µl water and an equal volume of 100% formamide loading dye was added. The RNA sample was boiled for 5 min and loaded on a denaturing acrylamide gel. The transcribed RNA was detected by UV-shadowing and purified according to the technique described in section 2.4.4.3.

*Production of the 3' half*: The transcription was performed in a 50 µl reaction and incubated for 2 h at 37°C. The reaction was prepared as follows:

| | |
|---|---|
| **5X Transcription buffer** | 10 µl |
| **DNA template** | 7 µg |
| **5X rNTPs mix** | 10 µl |

| | |
|---|---|
| **RNasin** | 2 µl |
| **T7 RNA polymerase** | 300 units |
| **50 mM 4-thioUpG or 50 mM UpG** | 4 µl |
| **Sterile water** | to a final volume of 50 µl |

- 5X Transcription buffer: 200 mM Tris-Cl pH 7.5, 10 mM spermidine, 50 mM DTT, 100 mM $MgCl_2$.

- 5X rNTPs mix: 5 mM each ATP, CTP and UTP, 2.5 mM GTP.

Typically one transcription reaction was prepared with 4-thioUpG and another one with UpG. The RNA produced with UpG was subsequently used as a control in UV-crosslinking experiments (section 2.7), the UpG being non-photoactivable should not give any UV-crosslinking signal.

The transcription reaction was stopped by adding 12.5 µl of 0.5 M EDTA. Phenol/chloroform (50 µl) was added to the reaction mix, after vortexing and spin in a microfuge, the aqueous phase containing the RNA was extracted and transferred to a fresh Ependorf tube. RNA was precipitated with ethanol. The RNA pellet was resuspended in 15 µl water and an equal volume of 100% formamide loading dye was added. The RNA sample was boiled for 5 min and loaded on a denaturing acrylamide gel. The RNA was visualized by UV-shadowing and purified according to the technique described in section 2.4.4.3.

### *5' labelling of the 3' half RNA moiety*

The 3' half-transcript was 5' phosphorylated in order to serve as a donor in the ligation reaction. The incorporation at this site of a $^{32}P$ allows incorporation of a single label within the RNA immediately adjacent of the site of crosslinking (the 4-thioU or the U residue).

The labelling reaction was prepared as follows and incubated for 30 min at 37°C:

| | |
|---|---|
| **50 pmole of 3'half transcript** | variable volume |
| **$^{32}$P-γATP (5000 Ci/mmole, 10 μCi/μl)** | 8 μl |
| **12 μM ATP** | 2 μl |
| **RNasin** | 1 μl |
| **10X pRNA cocktail** | 3 μl |
| **T4 Polynucleotide kinase** | 30 units |
| **Sterile water** | to a final volume of 30 μl |

- 10X pRNA cocktail: 500 mM Tris-Cl pH 8.0, 100 mM $MgCl_2$, 50 mM DTT.

2 μl of 0.5 M EDTA were added to stop the reaction and Phenol/chloroform (30 μl) was added to the reaction mix. After vortexing and spin in a microfuge, the aqueous phase containing the RNA, was extracted and transferred to a fresh Ependorf tube. RNA was precipitated with ethanol and pelleted by centrifugation at 4°C for 15 min.

### *Ligation of the two half-transcripts*

This ligation was mediated by a bridging DNA oligo complementary with the last 15-20 nucleotides of the 5' RNA half and with the first 15-20 nucleotides of the 3' RNA half (including the uridine or 4-thiouridine at the 5'extremity). The RNA ligation was mediated by a highly concentrated T4 DNA ligase. The protocol was applied as follows:

To the $^{32}$P-3'half RNA pellet, 100 pmole of the 5' half was added and the volume adjusted to 10 μl with sterile water. To this RNA solution were added:

| | |
|---|---|
| **10X pRNA cocktail** | 1.2 μl |
| **10 mM ATP** | 1 μl |
| **50 mM Bridging oligo** | 2 μl |

The mixture was boiled for 2 min at 90°C, then left to cool to room temperature over a period of 15 min. The following compounds were then added:

| | |
|---|---|
| **Ligase buffer** | 1 µl |
| **0.4 M DTT** | 1 µl |
| **RNasin** | 1 µl |
| **Highly concentrated T4 DNA ligase (2000 units/µl)** | 2 µl |

- 10X Ligase buffer: 500 mM Tris-Cl pH 8.0, 10 mM ATP, 10 mM $MgCl_2$, 1 µg/µl acetyl BSA.

The ligation reaction was incubated for 90 min at 37°C. An equal volume of 100% formamide loading dye was then added, the sample was boiled for 5 min and loaded on a denaturing acrylamide gel (section 2.4.4.2). The ligated product was detected by autoradiography (section 2.2.2) and purified according to the method described in section 2.4.4.3.

## 2.5 Proteins Methods

### 2.5.1 Quantification of total cellular protein

Protein concentrations were determined by the Bradford method using Bio-Rad protein detection solution as instructed by the manufacturer. BSA was used as standard.

### 2.5.2 Extraction of total cellular protein

***Small scale E. coli extracts***

*E. coli* cells were grown overnight at 37°C to saturation in the appropriate selective medium, typically LB-Amp (section 2.1.2.1). The following day, 50 µl of the overnight culture was used to inoculate 5 ml of fresh LB-Amp medium and the culture incubated at 37°C until the optical density at 550 nm reached 0.4. The culture was chilled on ice for 15 min, then 1 ml was transferred to an Eppendorf and the cells sedimented by centrifugation for 5 min at 4°C. The supernate was discarded, the cell

pellet resuspended in 100 µl of water and an equal volume of 2X SDS-sample buffer was added. The cells were lysed by heating to 100°C for 5 min. The lysate was centrifuged for 5 min to pellet cell debris, then 15-20 µl of the lysate loaded onto a SDS-polyacrylamide gel (section 2.5.3). The remaining supernate was frozen in dry ice and stored at -70°C.

### *Large scale E. coli extracts*

*E. coli* cells were grown overnight at 37°C to saturation in the appropriate selective medium, typically LB-Amp (section 2.1.2.1). The following day 100 µl of the overnight culture was used to inoculate 100 ml of fresh LB-Amp medium and the culture incubated at 37°C until the optical density at 550 nm reached 0.4. The culture was chilled on ice for 15 min then transferred to centrifugation bottle and centrifuged for 5 min at 4°C (Beckman JLA-25-10 rotor). The supernate was discarded, the cell pellet resuspended in 7 ml of lysis buffer and an equal volume of 2X SDS-sample buffer was added. The cells were lysed by heating to 100°C for 5 min. The lysate was centrifuged for 5 min to pellet cell debris. 15-20 µl of the lysate was loaded onto a 1 mm thick SDS-polyacrylamide gel (section 2.5.3) in order to check the protein content by coomassie staining (section 2.5.5). 0.5-1 ml of lysate was loaded on a 2 mm thick SDS-polyacrylamide gel for further purification of the protein of interest (section 2.10.2). The remaining supernate was frozen in dry ice and stored at -70°C.

- 2X SDS loading sample buffer: 125 mM Tris-Cl pH 6.8, 200 mM DTT, 4% (w/v) SDS, 40% (v/v) glycerol, 0.02% (w/v) Bromophenol blue dye.

- Lysis buffer: 50 mM Tris-Cl pH 7.5, 5 mM EDTA.

### *Yeast protein extracts*

Yeast cells were grown in 1 l of rich or selective medium (section 2.1.2.3) such that at the time of harvesting cells the optical density at 600 nm was between 0.6 and 2. The culture was harvested by centrifugation for 5 min at 3000 rpm and the supernate discarded. The cell pellet was treated according to the yeast splicing extract preparation protocol (section 2.6.1). After the homogenisation step, the homogenate was

centrifugated for 30 min at 17000 rpm (Rotor JLA-25 Beckman) (4°C). The supernate was harvested, aliquoted in Eppendorf tubes and snap frozen in dry ice. Samples were kept at -70°C for an indefinite time.

For SDS-polyacrylamide gel electrophoresis, protein extracts (typically 100 μg of total proteins) were mixed with an equal volume of 2X SDS-sample buffer (section 2.5.2) and the contents of the tube heated to 100°C for 5 min immediately prior to loading the gel.

### 2.5.3 SDS polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated as described by Laemmli (1970). All SDS-polyacrylamide gels were run using two 16 X 16 cm glass plates separated by either 1.5 or 2.0 mm spacers; 14 well combs accommodating up to 40 μl of sample, or single well combs for preparative gels were used. The resolving polyacrylamide gel solution was prepared following the recipe indicated in Table 5.1 and poured between the plates to a level of 4 cm from the top. The gel was layered with water-saturated butan-*2*-ol and allowed to polymerise at room temperature. After polymerisation the butan-*2*-ol was washed off with distilled water and a 4.8% polyacrylamide stacking gel (Table 11) poured on top of the resolving gel; the comb was inserted between the two plates, and the gel allowed to polymerise for at least 30 min at 37°C. Prior to electrophoresis the seal was removed from the plates, the comb gently taken out of the gel and the wells washed with distilled water. The plates were firmly fixed in the 'ATTO' electrophoresis apparatus and the chambers filled with 1X protein gel running buffer. Samples were loaded onto the gel beside pre-stained molecular weight markers, and the gel then run at 30 mA for the required time appropriate to the experiment or overnight at 8-10 mA.

- 10X Protein gel running buffer: 250 mM Tris-Base, 2.5 M glycine, 1% (w/v) SDS.
- High Molecular Weight pre-stained protein standards: Size range of 33-180 or 6.5-175 kD (Respectively obtained from Sigma and Biolabs).

◆ Low Molecular Weight pre-stained protein standards: Size range of 2.8-43 kD (Obtained from Gibco BRL).

**Table 11. SDS-Polyacrylamide gel electrophoresis recipe**

| Components | RESOLVING GEL | | | | STACKING GEL |
|---|---|---|---|---|---|
| | **6.0%** | **7.5%** | **8.5%** | **15%** | **4.8%** |
| **Water (ml)** | 16.5 | 15 | 14 | 7.5 | 5.9 |
| **30% Acrylamide (ml)** | 6.0 | 7.5 | 8.5 | 15.0 | 1.6 |
| **4X Gel Buffer (ml)** | 7.5 | 7.5 | 7.5 | 7.5 | 2.5 |
| **Fresh 10% (w/v) APS (µl)** | 55 | 55 | 55 | 120 | 30 |
| **TEMED (µl)** | 27.5 | 27.5 | 27.5 | 120 | 10 |

◆ 4X Resolving gel buffer: 1.5 M Tris-Cl pH 8.8, 0.4% (w/v) SDS.

◆ 4X Stacking gel buffer: 500 mM Tris-Cl pH 6.8, 0.4% (w/v) SDS.

### 2.5.4 Native gel for spliceosome complexes resolution

Two 16 X 16 cm glass plates separated by 1.5 mm spacers were taped together and prewarmed at 65°C. A blocking-gel was poured immediately and the gel-assembly left to stand at 65°C for 5 min. A separating-gel was then poured on top of the blocking-gel up to the top of the plates and a 10 x 1.0 cm comb was inserted. The gel was allowed to polymerise at room temperature for 30 min then stored at 4°C until use. Non-denaturing gels were run at 4°C in pre-chilled 1X native gel running buffer at 80 Volts for 10-12 h, or overnight at 40 Volts until the xylene blue dye front (section 2.6.2.2) had reached the blocking gel.

- Blocking Gel: 1 ml of 30% Acrylamide [30% (w/v) acrylamide:0.8% (w/v) bis-acrylamide], 100 µl of 0.5 M EDTA, 250 µl of 10X TB, sterile water to 5 ml. Then 50 µl of 10% (w/v) APS and 20 µl TEMED were added to allow gel polymerisation.

- Resolving Gel: 0.25 g agarose was dissolved in 40 ml sterile water and kept at 65°C. In another container, 5 ml of 30% Acrylamide [30% (w/v) acrylamide:0.8% (w/v) bis-acrylamide], 1 ml of 0.5 M EDTA, 2.5 ml of 10X TB were added. The dissolved agarose was added to this mixture and the volume adjusted to 50 ml. 400 µl of 10% (w/v) APS and 50 µl of TEMED was added and the gel poured immediately.

- 10X Native gel running buffer: 5X TB, 100 mM EDTA pH 8.0.

- 10X TB: 1.2 M Tris, 400 mM Boric acid.

### 2.5.5 Coomassie blue staining of polyacrylamide gels

100-500 ng of protein were detected by using this method. After electrophoresis the polyacrylamide gel was transferred to a clean plastic container and stained by addition of 5 gel volumes of Coomassie Blue staining solution. The gel was soaked in this solution for at least 30 min at room temperature with gentle shaking then the

staining solution was removed and the gel destained by incubating in [25% (v/v) methanol/10% (v/v) acetic acid] at room temperature. The destaining baths were regularly changed in order to accelerate the destaining process.

- Coomassie Blue staining solution: 0.25% (w/v) coomassie brilliant blue R250, 50% (v/v) methanol, 10% (v/v) acetic acid.

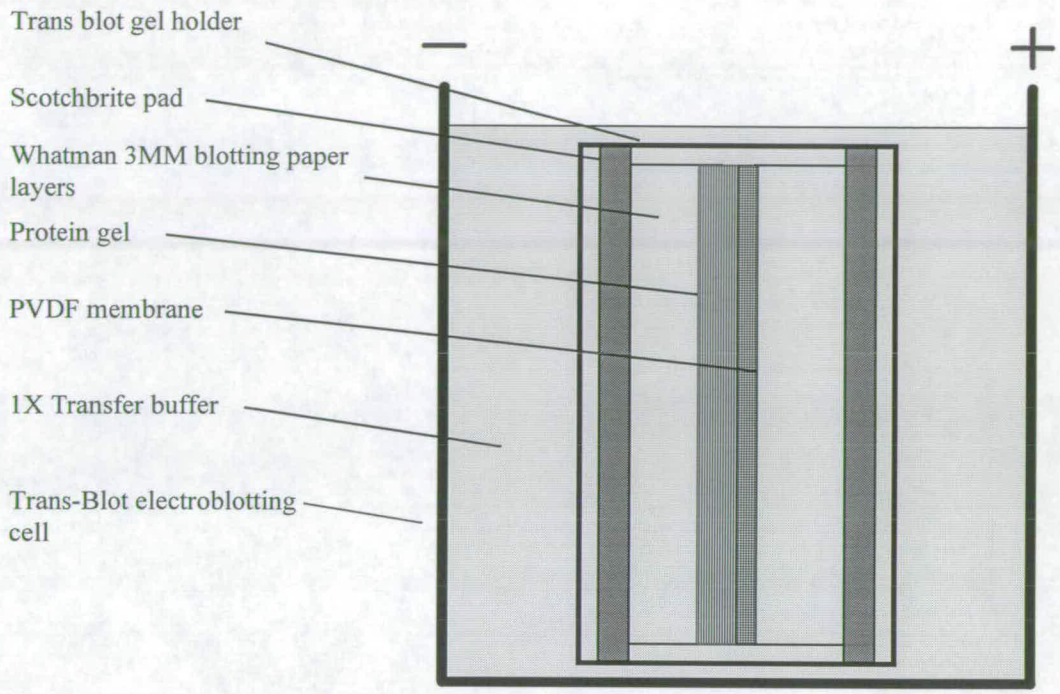
### 2.5.6 Western blotting

Following SDS-PAGE, proteins were electrophoretically transferred onto a polyvinylidene fluoride (PVDF) microporous membrane in a Trans-Blot Electroblotting Cell (Bio-Rad) with plate electrodes. Transferred proteins were visualized by Ponceau S (Sigma) staining, and the presence of the protein of interest detected using antibodies directed against this protein:

#### *Electrophoretic transfer of proteins onto PVDF Immobilon-P membrane*

Proteins were transferred electrophoretically onto Immobilon-P membrane (Millipore) using the Bio-Rad transfer system as instructed by the manufacturer. The gel was "sandwiched" as indicated in Figure 5.1, extreme care was taken to ensure that there were no air bubbles between the gel and membrane. Transfers were performed at 100 Volts using 1X western transfer buffer; for a time depending on the size of the proteins. For a 10-140 kD protein 30-60 min transfer was necessary; for a 140-220 kD protein 2-3 h were required.

**Figure 11. Sandwich for the protein transfer.**




- 10X Western transfer buffer: 1.5 M glycine, 200 mM Tris-Cl pH 8.3.

### *Blotting procedure*

Non-specific protein interactions were blocked by incubating the membrane in 5% (w/v) dry-milk or BSA (w/v) in 1X TBS (section 2.1.3) for 2 h at room temperature or overnight at 4°C. To this blotting solution, primary antibody was added at the appropriate dilution and incubated for further 2 h at room temperature. The membrane was then washed in 1X TBS for 15 min; this was repeated three times. Secondary antibody conjugated to alkaline phosphatase was diluted in 1X TBS/5% (w/v) dry-milk or BSA (w/v), added to the membrane and incubated for 1 h at room temperature. The membrane was washed thoroughly as before. Development of the membrane was achieved by exposing the membrane to alkaline phosphate substrates, NBT and BCIP according to the manufacturer's instructions.

## 2.6 *In vitro* splicing reaction

### 2.6.1 Yeast splicing extract preparation

The method described was adapted from R.J. Lin *et al* (1985). The splicing extract preparation required great care in order to avoid any RNase or protease contaminations, for this reason all the glass or plastic containers used were sterile. The dounce homogeniser used to homogenise cells was kept overnight in ethanol and then rinsed with water and finally with buffer A before use. A freshly grown yeast colony was used to inoculate the preculture, and a fresh preculture was used to inoculate the culture. The optical density (at 600 nm) at which the cells were collected varies from 1 to 2.4; for the SC261 yeast strain an $OD_{600}$=2.4 was preferred.

***The extract was prepared as follows***

Rich or selective medium (3 l) (section 2.1.2) were inoculated with a mid-logarithmic phase culture of the selected yeast strain and the culture grown overnight at 30°C. When the optical density at 600 nm was between 1 and 2.4 the cells were harvested by centrifugation for 5 min at 4000 rpm (Beckman JLA-10 500 rotor). The supernate was discarded, the cell pellet resuspended in 50 ml of 50 mM Potassium phosphate buffer pH 7.5 and transferred to a sterile 50 ml Falcon tube. Cells were pelleted by centrifugation for 5 min at 3000 rpm (Mistral 1000 centrifuge), the supernate discarded, and the pellet resuspended in 40 ml of lyticase buffer. The cell suspension was transferred to a 250 ml flask, 1 ml of a lyticase solution containing 2500 units enzyme dissolved in water was added, and cell suspension incubated at 30°C for 40-60 min with continuous slow shaking until 70-80% of the cells had formed spheroplasts, as determined microscopically. Spheroplasts were harvested by centrifugation for 5 min at 2000 rpm (Mistral 1000 centrifuge). The supernate was discarded and the spheroplasts then washed once with 1.2 M sorbitol and once with ice-cold SB-3 buffer; after each wash spheroplasts were gently resuspended using a sealed sterile pasteur pipette. After the final SB-3 wash the pellet was weighed, resuspended in approximately 1 ml/g of ice-cold Buffer A and transferred to a chilled Dounce

homogeniser. Spheroplasts were broken by 10-13 slow, rotating strokes of the tight fitting pestle; this was performed with the homogeniser maintained on ice. After homogenisation the lysate was transferred to a chilled, sterile 20 ml glass beaker, the volume of lysate measured and $1/9^{th}$ volume of cold 2 M KCl added dropwise while stirring with a magnetic stirrer. The homogenate was continuously stirred for 30 min at 4°C, then transferred to a chilled 15 ml polycarbonate tube and centrifuged for 15 min at 17000 rpm at 4°C in a pre-cooled Beckman JLA 25-10 rotor. Without disturbing the lipid layer, the supernate was then transferred to a chilled UltraPlus polycarbonate tube (Nalgene), and centrifuged at 38000 rpm for 70 min at 4°C in a pre-cooled Ti-70 rotor using a chilled Beckman XL-100 Ultracentrifuge. The final supernate was transferred to dialysis tubing (Visking size 1-8/32", 12-14 kD cut off) and dialysed against two changes of 1.5 l of ice cold Buffer D at 4°C over a period of 3 h. Following dialysis the extract was transferred to chilled Eppendorfs, centrifuged for 10 min at 4°C to remove insoluble matter, then aliquoted into fresh Eppendorfs and frozen in dry ice. Extracts were stored for long periods at -70°C; repeated freeze-thawing was avoided by storing the extract in small volumes.

- Lyticase buffer: 1.2 M Sorbitol, 50 mM Potassium phosphate buffer pH 7.5, 30 mM DTT.

- SB-3 Buffer$^{\Psi}$: 1.2 M Sorbitol, 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 3 mM DTT.

- Buffer A$^{\Psi}$: 10 mM HEPES-KOH pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT.

- Buffer D$^{\Psi}$: 20 mM HEPES-KOH pH 7.9, 0.2 mM EDTA, 0.5 mM DTT, 50 mM KCl, 20% (v/v) Glycerol.

All solutions were prepared in advanced (without DTT), sterilised by autoclaving. DTT was added to the solution just prior use.

$^{\Psi}$, prepared in advance and chilled overnight at 4°C.

***For the preparation of splicing extracts from galactose induced yeast cells***

Yeast cells were grown in YMGlyLacCas medium (section 2.1.2.3) containing the required amino acids for growth (Table 5). When the culture reached an $OD_{600}$=0.5-1.0, sterile galactose solution was added to a final concentration of 2% (w/v). The cells were grown for an extra 5-6 h then harvested and the splicing extract prepared as indicated above.

### 2.6.2 *In vitro* splicing assay

#### 2.6.2.1 *In vitro* splicing reaction and analysis of the splicing species on a denaturing gel

The assay for the production of spliced RNA and splicing intermediates was performed essentially as described by Newman *et al.* (1985). In a sterile Eppendorf tube the following components were added:

| | |
|---|---|
| **Splicing extract** | 4 µl |
| **5X Splicing buffer** | 2 µl |
| **30% (w/v) PEG 8000** | 1 µl |
| **$^{32}$P- pre-mRNA (20 000 cpm)** | 1 µl |
| **Sterile water** | 2 µl |

- 5X Splicing buffer: 12.5 mM $MgCl_2$, 10 mM ATP, 300 mM potassium phosphate buffer pH 7.5.

The mixture was incubated in a water bath at 24°C for 25 min, then stopped on ice.

For analysis of the RNA products, 6 µl of proteinase K solution was added to the splicing mix and the contents incubated at 37°C for 15 min. Following this incubation, 100 µl of splicing cocktail and an equal volume of phenol/chloroform were added.

After vortexing and centifugation, the aqueous phase containing the RNA was recovered and tranferred to a fresh Eppendorf tube. RNA was precipitated with ethanol.

The RNA pellet was centrifuged, dried and resuspended in 1.5 µl of water. 2 µl of 100% formamide loading dye was added, and the sample boiled for 5 min. The prepared RNAs were loaded on a 6% denaturing polyacrylamide gel to allow resolution of the splicing species which were subsequently visualised by autoradiography (section 2.2.2).

- Proteinase K solution: 2 mg/ml proteinase K, 50 mM EDTA, 1% (w/v) SDS.

- Splicing cocktail: 50 mM NaOAc pH 5.3, 1 mM EDTA, 0.1% (w/v) SDS, 25 µg/ml *E. coli* tRNA.

### 2.6.2.2 Analysis of spliceosome complexes on a native gel

This method was described by Pikielny *et al.* (1986). For the analysis of splicing complexes, the splicing reaction was stopped on ice, mixed with an equal volume of Q Buffer and incubated on ice for 10 min and then loaded directly onto a non-denaturing polyacrylamide gel (section 2.5.4) using non-denaturing gel-loading buffer.

- Q Buffer: 0.4 M KCl, 2 mM Magnesium acetate, 20 mM EDTA pH 8.0, 16 mM Tris-Cl pH 7.4, 2 mg/ml Yeast total RNA.

- 4X Non-denaturing gel loading dye: 50% (v/v) glycerol, 2.5% (v/v) 10X TB, 50 mM EDTA, 0.2% (w/v) bromophenol blue, 0.1% (w/v) xylene cyanol.

### 2.6.2.3 *In vitro* U6 snRNP reconstitution

The method used was adapted from Fabrizio *et al.* (1989). This is a two step reaction. During the first step the endogenous U6 snRNA was depleted due to endogenous RNase H activity of the splicing extract: A DNA oligo complementary to the U6 snRNA sequence 28-54 (Oligos V4528 and S3324, Table 8) was added to the splicing extract and the endogenous RNase H activated by incubating the extract at 30°C. This enzyme will then degrade all RNA involved in a RNA-DNA duplex. At this stage the splicing extract had lost its ability to splice due to the lack of U6 snRNA. The second step involves reconstitution of the U6 snRNP and the recovery of the splicing activity of the treated extract. The U6 snRNP was reconstituted by addition of *in vitro* transcribed U6 snRNA (section 2.4.4.1). $^{32}$P-radiolabelled pre-mRNA was also added in order to monitor the recovery of splicing acivity of the extract.

***Depletion of the endogenous U6 snRNA***

In a sterile eppendorf the following components were present and incubated for 35 min at 30°C:

| | |
|---|---|
| **Splicing extract** | 4 µl |
| **30% (w/v) PEG 8000** | 1 µl |
| **5X Splicing buffer** | 2 µl |
| **10X Knockout oligo** | 1 µl |
| **Sterile water** | 1 µl |

After incubation, sample were placed on ice.

- Splicing buffer: 12.5 mM $MgCl_2$, 10 mM ATP, 300 mM Potassium phosphate buffer pH 7.5.

- 10X Knockout oligo: The concentration of this oligo stock varies from one extract to another. Preliminary experiments were done for each extract in order to determine the minimum amount of this oligo needed to deplete the endogenous U6 snRNA and

abolish the splicing activity of the extract. Typically a concentration of 1.1-1.4 μM of this 10X oligo stock solution was required.

### *U6 snRNP reconstitution*

The U6 snRNA produced by transcription *in vitro* was added at the same time as the radiolabelled substrate. The following components were added to the cooled sample and incubated for 25 min at 24°C:

| | |
|---|---|
| ***In vitro* transcribed U6 snRNA** | 1μl |
| **$^{32}$P-pre-mRNA (20000 cpm)** | 1μl |

- The quantity of *in vitro* transcribed U6 snRNA added in the reconstitution reaction varies from 20 fmole to 1 pmole. This snRNA can be either unlabelled or $^{32}$P-radiolabelled.

The reaction was stopped by placing the sample on ice. For the analysis of the splicing species, the sample was treated as indicated in section 2.6.2.1. For the detection of RNA-Protein interactions the sample was then submitted to UV irradiation as indicated in section 2.7. Depending on the purposes of the study, an immunoprecipitation was carried out (section 2.8).

### 2.6.2.4 Depletion of the endogenous U5 snRNA in a yeast splicing extract

The depletion was done as described by O'Keefe *et al.* (1996). Splicing extracts were prepared from the Rok8 strain (Table 2.4) which contains as the only source of U5 snRNA a modified U5 snRNA containing 30 extra nucleotides inserted between nucleotides 85 and 86. With this additional sequence the modified U5 snRNA was then succeptible to RNase H degradation using knockout oligos Rok8L and Rok8L (Table 2.5). The depletion reaction was prepared as follows:

| | |
|---|---|
| **Rok8 Splicing extract** | 4.8 µl |
| **5X Splicing buffer** | 2 µl |
| **15X Rok8L and Rok8R Knockout oligos** | 0.7 µl |
| **Water** | To final volume of 10 µl |

The mix was incubated at 34°C for 30 min then placed on ice.

- 15X Rok8L and Rok8R knockout oligo: each oligo was present at a concentration of 700 nM.

#### 2.6.2.5 Depletion of the endogenous U2 snRNA in a yeast splicing extract

The depletion was done as described by McPheeters *et al.* (1989). Using the same principle and recipe as for the endogenous U6 snRNA depletion, a U2 snRNA knockout oligo SRU2 (Table 8) was added to the yeast splicing extract. SRU2 was present at a final concentration of 600 nM. In order to completely degrade the endogenous U2 snRNA, the amount of SRU2 oligo added is in excess compared to the minimal quantity required to deplete the endogenous U2 snRNA.

## 2.7 Protein-RNA UV-Crosslinking

The sample to be crosslinked was placed in a pre-cooled microtiter plate (Greiner). 1µl of a 10 mg/ml *E. coli* t-RNA stock was added, the plate directly positioned under the UV-lamp bulb and kept on ice.

***For the UV-crosslinking of protein to $^{32}$P-RNA***: Samples were exposed to a UV light with a wavelength of 254 nm for 25 min (Stratagene UV Stratalinker™ ).

***For the UV-crosslinking of protein to $^{32}$P-RNA containing a photoactivatable 4-thiouridine***: Samples were exposed to a UV light with a wavelength of 365 nm for 5 min (UVP,inc; Model B-100AP).

After the irradiation, samples were transferred to pre-cooled sterile Eppendorf tubes and were ready for subsequent experiments.

## 2.8 Immunoprecipitation

Antibodies (Table 10) were bound to pre-swollen protein A-Sepharose beads (8 mg PAS per immunoprecipitation sample; Sigma) in NTN buffer for 2 h at room temperature. To PAS-bound antibodies, 200 µl of blocking solution was added and samples incubated for a further 1 h at room temperature. The PAS-bound antibodies were washed four times with NTN buffer and ready for use.

To the sample to be immunoprecipitated, an equal volume of 2X IQ buffer was added and the volume was adjusted (if necessary) to 300 µl with 1X IQ buffer. Samples were incubated at 4°C with PAS-bound antibodies for 2 h with constant rotation. The antibody complexes were then washed four times with NTN, and once with NT (NTN without NP-40); the salt concentration of NTN was varied according to the stringency of the wash required. Depending on the purpose of the immunoprecipitation, the immunoprecipitate was treated with different methods:

### 2.8.1 Immunoprecipitation of proteins

Immunoprecipitated proteins were recovered from the PAS-bound antibodies by resuspending the PAS-matrix in 50 µl of 2X SDS-sample buffer (section 2.5.2), heating the sample to 100°C for 5 min then pelleting the matrix by centrifugation for 5 min. The

protein-containing supernate was resolved on an SDS-polyacrylamide gel (section 2.5.3).

### 2.8.2 Immunoprecipitation of snRNAs

Co-immunoprecipitated snRNAs were recovered from the PAS-bound antibodies by treatment with 50 µl of 2 mg/ml proteinase K (IP grade) for 30 min at 37°C. Sterile water (100 µl) was added to the Eppendorf and the resulting mixture extracted once with an equal volume of phenol:chloroform. Approximately 200 µl of the aqueous phase was transferred to a fresh Eppendorf, a further 100 µl of splicing cocktail (section 2.7.2.1) was then added to the remaining solution and the extraction repeated. 1 µl of 10 mg/ml *E. coli* tRNA solution and 0.1 volume of 3 M sodium acetate (pH 5.2) were added to the pooled aqueous phases and the RNA recovered by ethanol precipitation. The RNA pellet was dried, resuspended in 5 µl of sterile water and 5 µl of 100% formamide loading dye and then fractionated on a denaturing polyacrylamide gel (section 2.4.4.1).

- NTN: 150 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.1% (v/v) NP-40.

- Blocking solution: 100 µg/ml BSA, 100 µg/ml glycogen, 100 µg/ml *E.coli* tRNA.

- 2X IQ buffer: 300 mM NaCl, 5 mM $MgCl_2$, 15 mM Hepes pH 7.5, 0.1% (v/v) NP-40.

- Proteinase K (IP Grade): 50 mM Tris-Cl pH 7.5, 300 mM NaCl, 5 mM EDTA, 1.5% (w/v) SDS, 2 mg/ml proteinase K.

## 2.9 Immunisation Procedure

### 2.9.1 Expression of *LacZ*, *trpE* and *GST* fusion proteins

Three expression systems were used to produce large amounts of fusion protein in bacteria. *LacZ*-gene fusions were generated using the pUR series of vectors (Rüther and Müller-Hill, 1983), *trpE*-gene fusions were constructed using pATH vectors (Koerner *et al.*, 1990) and finally GST-gene fusions were constructed using pGEX-2T vector (Pharmacia). The gene of interest was subcloned into either a pUR, pATH or pGEX-2T vector to create an in-frame fusion with respectively the *lacZ*, *trpE* or GST open reading frame (ORF) (Table 9)

#### 2.9.1.1 Induction of lacZ fusion proteins

*E. coli* cells (BMH71-18; Table 6) carrying the desired *lacZ*-gene fusion were grown overnight to saturation at 37°C in LB-Amp medium. The following morning, the culture was diluted by a factor of ten with fresh, pre-warmed LB-medium and the incubation continued with vigorous shaking at 37°C until $OD_{550}$=0.4-0.7. IPTG was added to the culture to a final concentration of 0.2 mM-1 mM and the culture then incubated at 37°C for a further 2-4 h.

#### 2.9.1.2 Induction of trpE fusion proteins

Induction of *trpE*-fusion protein synthesis was performed as described by Koerner *et al.* (1990). *E. coli* cells (DH5αF'; Table 6) carrying the desired *trpE*-gene fusion were grown overnight to saturation at 37°C in synthetic M9 medium supplemented with 20 μg/ml tryptophan. The following morning, the culture was diluted by a factor of ten with fresh, pre-warmed synthetic M9 medium lacking tryptophan, and incubation continued with vigorous shaking at 37°C for 1 h.

Indoleacrylic acid (IAA; 2 mg/ml in 95% ethanol) was added to a final concentration of 0.01 mg/ml and the culture incubated at 37°C for a further 4 h.

### 2.9.1.3 Induction of GST-fusion proteins

Induction of GST-fusion protein synthesis was performed according to the manufacturer's instructions (Pharmacia). *E. coli* cells (DH5αF'; Table 6) carrying the desired GST-gene fusion were grown overnight to saturation at 37°C in LB-Amp medium (Section 2.1.2.1). The following morning, the culture was diluted by a factor of ten with fresh, pre-warmed LB-Amp medium, and incubation continued with vigorous shaking at 37°C until $OD_{600}$=1-2. IPTG was added to a final concentration of 1 mM and the culture incubated at 37°C for additional 2-6 h.

### 2.9.1.4 Induction of 6X histidine-fusion proteins

Induction of 6X His-fusion protein synthesis was performed according to the manufacturer's instructions (Qiagen). *E. coli* cells (BMH71-18; Table 6) carrying the desired 6X His-gene fusion were grown overnight to saturation at 37°C in LB-Amp medium (Section 2.1.2.1). The following morning, the culture was diluted by a factor of ten with fresh, pre-warmed LB-Amp medium, and incubation continued with vigorous shaking at 37°C until $OD_{600}$=1-2. IPTG was added to a final concentration of 1 mM and the culture incubated at 37°C for additional 2-6 h.

The amount of inducing compound required to induce a high fusion protein expression varies. Preliminary experiments were done testing different amounts of inducing chemicals, induction times and growth temperatures in order to optimize the induction conditions.

Induced cultures were harvested by centrifugation and processed according to the method described in section 2.5.2, depending on the purposes of the experiment.

### 2.9.2 Partial purification of fusion protein

Large scale protein extracts prepared from *E.coli* cells overproducing fusion protein were fractionated on a preparative 2 mm thick 8% SDS-polyacrylamide gel (sections 2.5.2 and 2.5.3). The protein sample was run beside a pre-stained molecular weight and the gel run such as a good resolution of the fusion protein was obtained. Following electrophoresis, 0.5 cm strips were cut from either side of the polyacrylamide gel and briefly stained with Coomassie Blue (section 2.5.5) to locate the fusion protein, the remainder of the gel was kept at 4°C wrapped in a plastic film with 3 ml of 1X protein gel running buffer. The stained strips were aligned with the remainder of the polyacrylamide gel and the band of interest excised. This gel slice was placed in dialysis tubing (Visking size 1-8/32", 12-14 kD cut off) and filled with 2-3 ml of 1X protein gel running buffer (section 2.5.2). The dialysis tubing was sealed, placed crosswise in a horizontal electrophoresis chamber containing 1X protein gel running buffer, and the protein electroeluted from the gel slice for 3-4 h at 100 V. Following electroelution, the current was reversed for 1 min and the gel buffer removed from the dialysis tubing. The protein-containing buffer was placed in the sample reservoir of a Centricon centrifugal concentrator and concentrated according to the manufacturers' instructions. The concentrated fusion protein was aliquoted in sterile Eppendorf tubes and stored at -70°C. The fusion protein concentration was estimated by running several dilutions of the concentrate beside known quantities of BSA on an SDS-polyacrylamide gel. The gel was then Coomassie stained (section 2.5.2) and the fusion-protein concentration evaluated by comparison with the known BSA amounts present on the same gel.

### 2.9.3 Immunisation of rabbits

All experiments on live animals were performed in compliance with the "Cruelty to Animals Act" 1986 under licence 63181. New Zealand White female rabbits were used.

#### 2.9.3.1 Immunisation

For the initial immunisation, 150 μg of partially purified fusion protein (section 2.9.2) was diluted into sterile water to give a final volume of 500 μl. An equal volume of Freund's Complete adjuvant (Sigma) was added and the compounds mixed by using a 2 ml syringe: an emulsion forms by vigorous, prolonged mixing. This was injected into the rabbit subcutaneously. All subsequent injections consisted of 100 μg of fusion protein mixed with Freund's Incomplete adjuvant. Immunisations were performed at 4-6 weekly intervals, and a test bleed done 10 days after the injection.

#### 2.9.3.2 Bleeds

Blood was collected from the rabbit prior to the first immunisation (pre-immune) and 10-12 days after the fourth and all subsequent injections. 10 days after the final immunisation, when the immune response had reached a peak, the rabbit was anaesthetised with Sagatal, 100-150 ml of blood withdrawn directly from the heart and the rabbit killed with a further injection of Sagatal.

### 2.9.4 Serum preparation and storage

After collection the blood was allowed to clot at 4°C overnight. The clot was then detached from the side of the container using a sterile Pasteur pipette, and serum aspirated from the container with the pasteur pipette avoiding disturbing the clot. The serum was separated from blood cells by centrifugation for 5 min at 2000 rpm, transferred to a fresh tube and aliquoted into Eppendorf tubes stored at -70°C.

### 2.9.5 Affinity purification of antibodies

Crude lysate from *E. coli* overexpressing a fusion protein detected by the antibodies to be purified was loaded on a SDS polyacrylamide gel (section 2.5.3). The proteins were transferred onto a PVDF-immobilon-P membrane according to the transfer method described in section 2.5.5 and stained by ponceau staining (Sigma) according to the manufacturers' instructions. The band corresponding to the fusion protein was cut out and the membrane incubated for 2-4 h at room temperature in 1X TBS/1% (w/v) BSA or 1% (w/v) low fat milk powder. After this incubation several dilutions of crude serum (between 1/100 and 1/300 dilution) were added to blotted membrane strips and left to bind to the membranes for 2-3 h at room temperature with gentle shaking. The strips were washed five time with 1X TBS/0.1% (v/v) NP-40 solution (section 2.1.3) over a period of 1 h. Antibodies were eluted by incubating the membrane with 2 ml of elution solution for 5 min, the eluate was recovered and immediately neutralized by the addition of 0.5 ml of 1 M Tris-Cl pH 7.5. A second elution was done using 2 ml of the elution solution and left to incubate for 10 min, then the eluate was collected and neutralised as indicated previously. The two eluates were pooled and dialysed against three changes of 1 l each of 1X TBS (section 2.1.3) at 4°C over a period of 16-20 h. Affinity purified antibodies were stored at 4°C in the presence of 0.1% (w/v) sodium azide.

- Elution solution: 0.1 M glycine-HCL pH 2.5, 0.1% (w/v) BSA.

### 2.9.6 Covalently coupling of antibodies to the Affi-Gel HZ matrix

Affi-Gel HZ is an agarose support (purchased from Bio-Rad) which reacts with the aldehydes of oxidized carbohydrates to form stable, covalent hydrazone bonds. In order to covalently attach antibodies to the matrix, the heavy chain of the immunoglobulin G is oxidized with periodate. The antibodies are therefore attached to the agarose gel matrix by their Fc region which orientate the coupling to the support and improve the binding capacity of the Ab-agarose gel. The preparation of the

antibodies and the coupling to the Affi gel HZ were carried out according to the manufacturers' instructions.

# CHAPTER THREE

## Preparation of Anti-Prp24 Antibodies as a Tool to Investigate Protein-U6 SnRNA Interactions in the U6 SnRNP

## INTRODUCTION

*prp24-1* was originally identified as a temperature-sensitive yeast mutant which accumulates pre-mRNA at the non-permissive temperature (Vijayraghavan *et al.*, 1989). Mutations in the *PRP24* gene were shown to suppress a cold-sensitive phenotype associated with the *snr14-G14C* mutation (Shannon and Guthrie, 1991). The *snr14-G14C* mutant contains a G to C change at U4 snRNA position 14 within the region of interaction with U6 snRNA (stem II) which affects U4/U6 snRNP stability. Further analysis indicated that Prp24p possesses three RNA-binding domains and it directly contacts the U6 snRNA in the U6 snRNP particle (Shannon *et al.*, 1991; Jandrositz and Guthrie, 1995). *In vitro* binding assays demonstrated that Prp24p purified from *E. coli* can bind to the U6 snRNA at nucleotides 39-56 and to the U6 snRNA at nucleotides 39-56 and 64-76 when U6 snRNA is base-paired with the U4 snRNA (Ghetti *et al.*, 1995). In yeast Prp24p interacts transiently with the U4 snRNA and the U6 snRNA promoting U4/U6 snRNP formation but it must leave this complex to allow completion of U4/U6 snRNA interaction (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995). In addition, Prp24p in the free U6 snRNP promotes the reannealing of the U6 snRNA with the U4 snRNA freed from the disassembled spliceosomes, recycling the U4/U6 snRNP (Raghunathan and Guthrie, 1998).

As an internal positive control in our UV-crosslinking studies, we wanted to show using anti-Prp24 antibodies that Prp24p contacted U6 snRNA. As no anti-Prp24 antiserum was available in our laboratory, Prp24p was tagged with a FLAG epitope. The description of the cloning and the functional analysis of epitope-tagged Prp24p will be presented in this chapter. The generation of polyclonal anti-Prp24 antisera by immunising rabbits with Prp24p expressed in and purified from *E.coli* will be described.

## 3.1 Disruption of the *PRP24* Gene in *Saccharomyces cerevisiae*

The *PRP24* gene was disrupted according to the method described by Baudin *et al.* (1993) (Figure 12, Panel A). For this purpose a disrupted version of the *PRP24* gene was constructed which contains the *HIS3* selectable marker and 34 and 28 nucleotides of homology with the 5' and the 3' ends respectively of *PRP24* ORF. The *HIS3* ORF was inserted in the opposite orientation to the *PRP24* ORF to avoid any potential fusion with the remaining Prp24p-encoding sequence. Diploid yeast cells (DJY84; Table 7) were transformed (section 2.3.3) with the disrupted construct and selected for their histidine prototrophy. Yeast colonies containing a disrupted genomic copy of *PRP24* were identified by PCR (section 2.4.3.8) with oligodeoxynucleotides P5118 and P5117 (Table 8) used as 5' and 3' primers respectively. A 982 bp DNA fragment was expected from the amplification of the wild-type copy of *PRP24* while a 1030 bp DNA fragment was expected for the amplification of the disrupted copy. All nine transformants tested gave the correct PCR products.

One diploid (YVDP24; Table 7) harboring a disrupted copy of *PRP24* was sporulated and twenty tetrads were dissected (section 2.3.4). Only two spores were viable, and neither of these contained the *HIS3* marker used for the disruption, demonstrating that a lethal phenotype was associated with *PRP24* gene disruption.

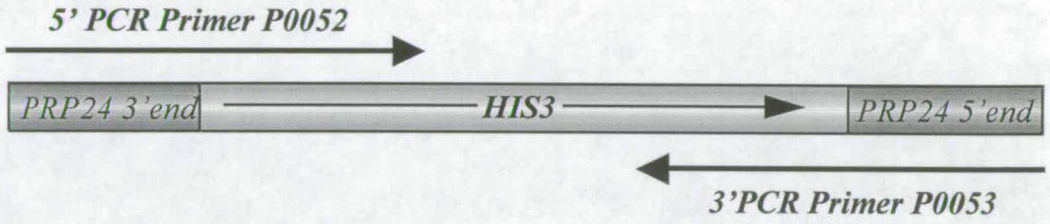
*PRP24* gene disruption was confirmed by Southern blotting analysis (section 2.4.3.11) (Figure 12; Panel B). The genome analysis of haploid cells obtained from the germination of the two viable spores demonstrated that both possessed a wild-type copy of the *PRP24* gene.

**Figure 12. Disruption of the *PRP24* gene in yeast.**

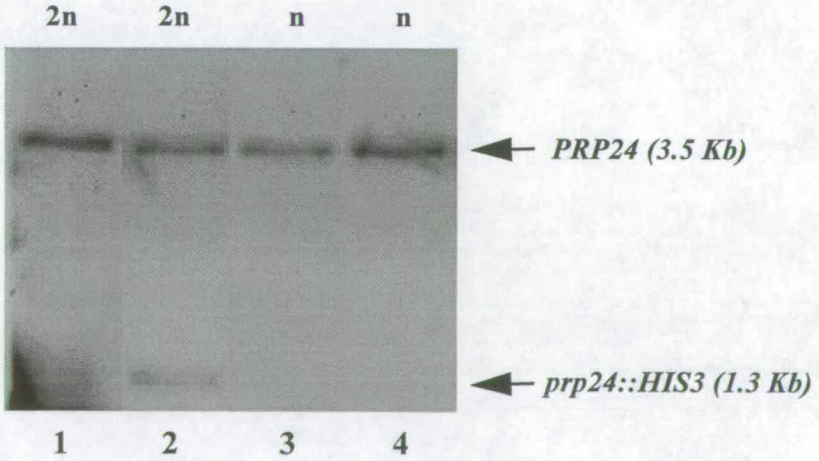
**A:** Construction of a disrupted *prp24::HIS3* allele. This construct was generated by PCR (section 2.4.3.8) carried out on YIP1 plasmid DNA containing the *HIS3* gene (Table 9). Oligodeoxynucleotides P0052 and P0053 (Table 8) were used as 5' and 3' primers respectively. The amplified DNA fragment possesses the *HIS3* sequence and 34 nt and 28 nt homology with the 5' and the 3' ends respectively of *PRP24* ORF.

**B:** Southern blotting analysis. Genomic DNA was prepared (section 2.4.3.2) from diploid yeast strain DJY84 (Table 7) (lane 1) and strain YVDP24 (shown by PCR analysis to contain a disrupted genomic copy of *PRP24*) (lane 2). Genomic DNA was also prepared from the two viable haploid cells obtained after sporulation of YVDP24 and tetrad dissection (lanes 3 and 4). Genomic DNA preparations were completely digested with *Xba*I and *Bgl*I restriction enzymes and DNA fragments were resolved on an agarose gel then transferred to a Nylon membrane (section 2.4.4.3). The probe used to detect *PRP24* gene was amplified by PCR from YCpXba plasmid (Table 9) using oligodeoxynucleotides JMB10 and JMB11 (Table 8) as primers. This probe annealed to the last 369 bp of *PRP24* ORF and to 396 bp of the downstream sequence and was radiolabelled by random priming (section 2.4.3.9). A DNA fragment of 3.5 kb was expected for the detection of the wild-type *PRP24* gene, while a 1.3 kb fragment was expected for the detection of the disrupted allele.

A




B




**Southern blot analysis**

## 3.2 Cloning of the Gene Encoding the FLAG Epitope-Prp24 Fusion Protein

Prp24p was tagged with the FLAG epitope (Kodak, IBI FLAG system) which is a highly hydrophilic octapeptide (N terminus-D-Y-K-D-D-D-D-K-) against which mouse monoclonal antibodies are commercially available (Kodak). It was hoped that the insertion of the FLAG epitope at the C-terminus of Prp24p would not interfere with its function. The cloning strategy to fuse the FLAG-epitope DNA sequence to the *PRP24* gene is described in Figure 13. The *PRP24* gene sequence was communicated by A. Ghetti (laboratory of J. Abelson, Cal. Tech) (Annex I), and the YCpXba plasmid (*CEN-URA3-PRP24*; Table 9) was provided by P. Raghunathan (laboratory of C. Guthrie, UCSF). YCpXba contains a 3.5 kb insert including the *PRP24* ORF and flanking sequences (Shannon *et al.*, 1991). Sequence analysis indicated that the fusion between the *PRP24* ORF and the FLAG encoding sequence in the YRFL24 construct (Figure 13) was correct. The functionality of the tagged Prp24p was tested by complementation of a lethal phenotype associated with *PRP24* disruption. For this purpose, the diploid yeast strain YVDP24 (*prp24::HIS3/PRP24*; Table 7) was transformed (section 2.3.3) with the plasmid YRFL24. The selected transformants were sporulated (section 2.3.4), tetrads dissected and grown on YPDA agar (section 2.1.2.3). For each tetrad dissected, only two spores grew out of four, the analysis of their genotype showed that the viable cells contained a wild-type copy of the *PRP24* gene. No rescue of the *PRP24* null allele was obtained with the plasmid YRFL24. Several explanations can be proposed:

1) No transcription termination sequence was present at the 3' end of *PRP24-FLAG* ORF. Therefore the mRNA may be incorrectly processed at its 3' end and consequently made unstable. This would result in fusion protein not being produced.

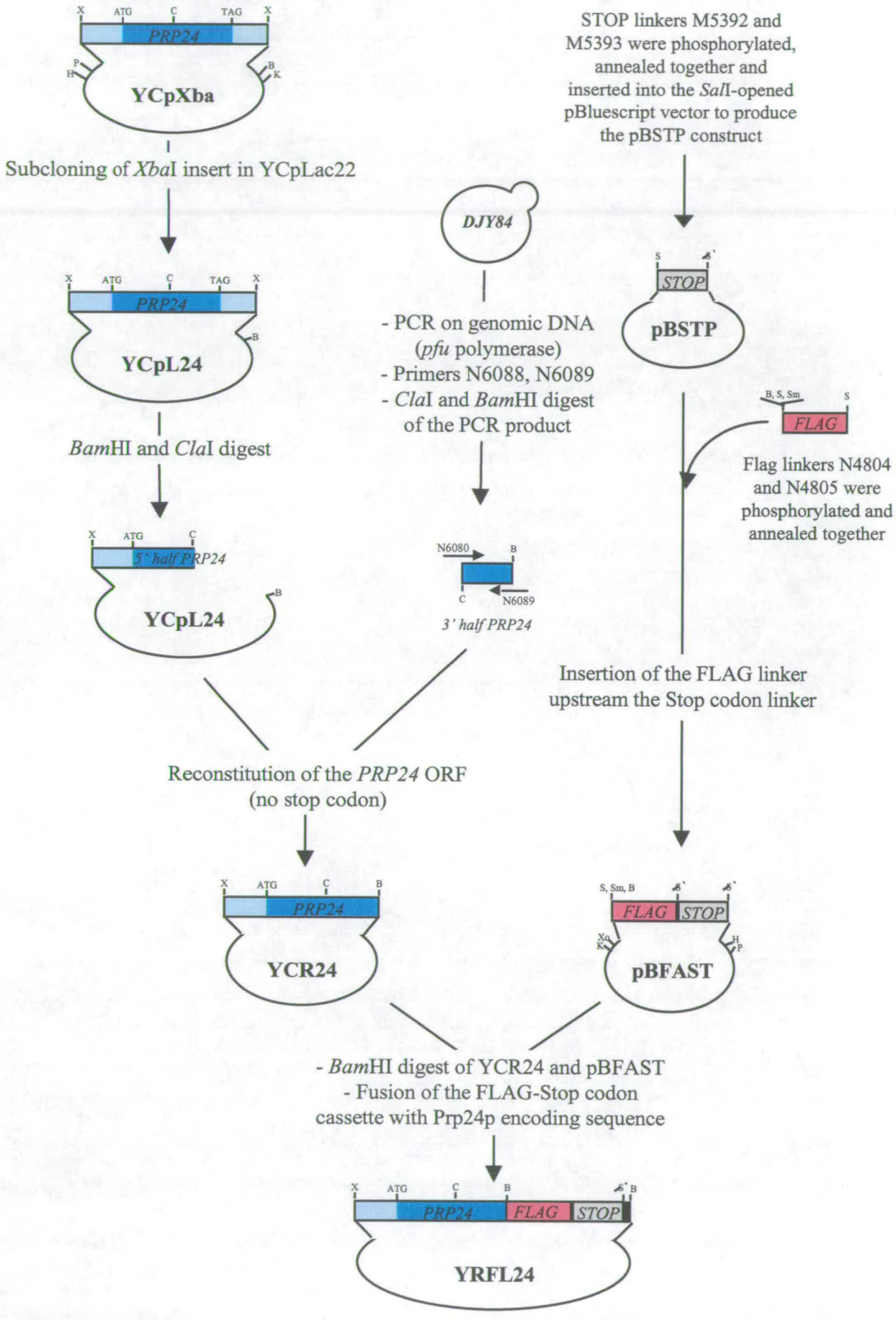
2) Although the sequence of the *PRP24-FLAG* junction was correct, the high fidelity *pfu* polymerase used to amplify the 3' half of *PRP24* ORF (Figure 13) may have inserted mutation(s) responsible for the lack of function of the epitope tagged Prp24p.

3) The addition of 15 extra amino acids at the C-terminus of Prp24p may have affected the function of the fusion protein. Raghunathan *et al.* (1997), have inserted at the C-

**Figure 13. Cloning strategy followed to fuse the FLAG epitope-encoding sequence downstream of the *PRP24* ORF.**

The fusion protein is tagged at its C-terminus with 15 amino acids.

B, *Bam*HI; C, *Cla*I; H, *Hind*III; K, *Knp*I; P, *Pst*I; S, *Sal*I; crossed S, lost *Sal*I site; Sm, *Sma*I; X, *Xba*I; Xo, *Xho*I.


X
ATG
C
TAG
X
PRP24
P
H
B
K
YCpXba
Subcloning of XbaI insert in YCpLac22
X
ATG
C
TAG
X
PRP24
B
YCpL24
BamHI and ClaI digest
X
ATG
C
5' half PRP24
B
YCpL24
DJY84
- PCR on genomic DNA
(pfu polymerase)
- Primers N6088, N6089
- ClaI and BamHI digest
of the PCR product
N6080
B
C
N6089
3' half PRP24
STOP linkers M5392 and
M5393 were phosphorylated,
annealed together and
inserted into the SalI-opened
pBluescript vector to produce
the pBSTP construct
S
STOP
pBSTP
B, S, Sm
S
FLAG
Flag linkers N4804
and N4805 were
phosphorylated and
annealed together
Insertion of the FLAG linker
upstream the Stop codon linker
Reconstitution of the PRP24 ORF
(no stop codon)
X
ATG
C
B
PRP24
YCR24
S, Sm, B
FLAG
STOP
Xo
K
H
P
pBFAST
- BamHI digest of YCR24 and pBFAST
- Fusion of the FLAG-Stop codon
cassette with Prp24p encoding sequence
X
ATG
C
B
PRP24
FLAG
STOP
B
YRFL24

terminus of Prp24p three Hemagglutinin (HA) epitopes (27 amino acids) and found that although the fusion protein was functional, it caused a detrimental effect on cell growth and on the stability of the U6 snRNA (Raghunathan, personnal communication). These data suggest that Prp24p may not tolerate very well extra amino acids at its C-terminus. If the lack of function of the Prp24-FLAG fusion protein was due to the presence of the FLAG epitope, the chemical nature of the 15 amino acids added rather than their number was more likely to be reponsible for this dramatic effect.

## 3.3 Production of Anti-Prp24 Antisera from Rabbits

### 3.3.1 Expression of the *PRP24* gene in *E.coli*

Four different expression systems were tested for the producion of high amounts of Prp24p in *E. coli* cells. For this purpose three plasmids (pATP24, pGEX24, pGAP24, pIP24) derived from pATH3, pGEX-2T, pUR288 and pQE-30 respectively were constructed and the cloning strategy followed is described in Figure 14.

***Production of anthranilate synthase-Prp24 fusion protein with pATP24:*** the sequence analysis showed that the *PRP24* ORF amplified by PCR from yeast genomic DNA and cloned into the pATH3 vector was identical to that contained in the plasmid YCpXba (Table 9). *DH5αF'* (pATP24) *E. coli* cells (Table 6) were induced for the production of the anthranilate-Prp24 fusion protein by addition of indoleacrylic acid (IAA) (section 2.9.1). The production of the fusion protein (87 kD) was confirmed by Coomassie blue staining of proteins after fractionation on an SDS-PAGE gel (data not shown). In addition to the trpE-Prp24 protein, numerous unrelated polypeptides of similar size were concomitantly produced and interfered with the partial purification of the fusion protein (section 2.9.2), preventing its utilization to raise antibodies from rabbits.

**Figure 14. Cloning strategy followed to construct pATP24, pGEX24, pGAP24 and pIP24 plasmids.**

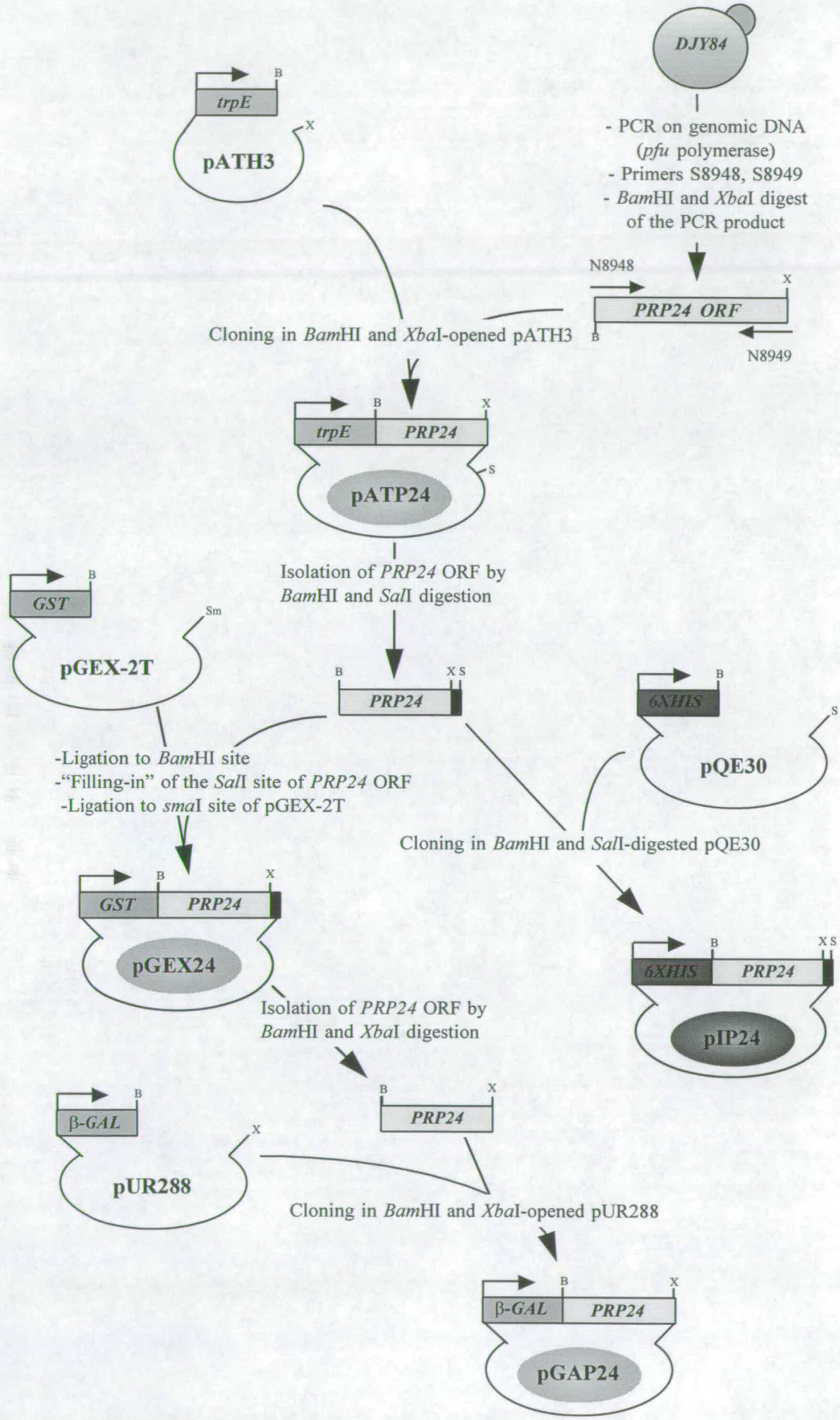
For pATP24: the *PRP24* ORF was amplified by PCR with the *pfu* polymerase using genomic yeast DNA prepared from the wild-type DJY84 strain (Table 7) as template. Oligodeoxynucleotides S8948 and S8949 (Table 8) were utilized as primers. The resulting PCR product possessed the entire *PRP24* ORF lacking any flanking sequence and two restriction sites were engineered at each extremity (*Bam*HI at the 5' end and *Xba*I at the 3' end). After *Bam*HI and *Xba*I digestion the amplified DNA was cloned in frame between the *Bam*HI and *Xba*I restriction sites of the *trpE*-expression vector (pATH3; Koerner *et al.*, 1990). The *PRP24* ORF was inserted in frame with the anthranylate synthase encoding gene.

**For pGEX24:** the *PRP24* ORF was isolated from pATP24 vector by *Bam*HI and *Sal*I digestion and was ligated to the *Bam*HI end of the *Bam*HI-*Sma*I opened-pGEM-2T vector. Then the *Sal*I extremity of the *PRP24* ORF was filled-in by the klenow enzyme and ligated to the *Sma*I end of the vector. *PRP24* ORF was cloned in frame with the 3' end of the GST encoding sequence in the pGEX-2T expression vector (Pharmacia).

**For pIP24:** *PRP24* ORF was isolated from pATP24 after digestion with *Bam*HI and *Sal*I and ligated to the *Bam*HI/*Sal*I-opened pQE-30 vector (Qiagen). The resulting pIP24 construct harbors a sequence encoding for 6 repeats of histidine fused in frame with PRP24 encoding sequence.

**For pGAP24:** the *PRP24* ORF was isolated from pGEX24 by *Bam*HI and *Xba*I digestion and was inserted between the *Bam*HI and *Xba*I restriction sites of the expression vector pUR288 (Rüther *et al.*, 1983). The resulting pGAP24 vector possesses the Prp24p encoding sequence fused in frame to the β-*galactosidase* gene.

B: *Bam*HI; S: *Sal*I; Sm: *Sma*I; X: *Xba*I.


DJY84
- PCR on genomic DNA
(*pfu* polymerase)
- Primers S8948, S8949
- *Bam*HI and *Xba*I digest
of the PCR product
N8948
N8949
PRP24 ORF
trpE
pATH3
Cloning in *Bam*HI and *Xba*I-opened pATH3
trpE
PRP24
pATP24
Isolation of *PRP24* ORF by
*Bam*HI and *Sal*I digestion
GST
pGEX-2T
PRP24
6XHIS
pQE30
-Ligation to *Bam*HI site
-"Filling-in" of the *Sal*I site of *PRP24* ORF
-Ligation to *sma*I site of pGEX-2T
Cloning in *Bam*HI and *Sal*I-digested pQE30
GST
PRP24
pGEX24
6XHIS
PRP24
pIP24
Isolation of *PRP24* ORF by
*Bam*HI and *Xba*I digestion
β-GAL
pUR288
PRP24
Cloning in *Bam*HI and *Xba*I-opened pUR288
β-GAL
PRP24
pGAP24

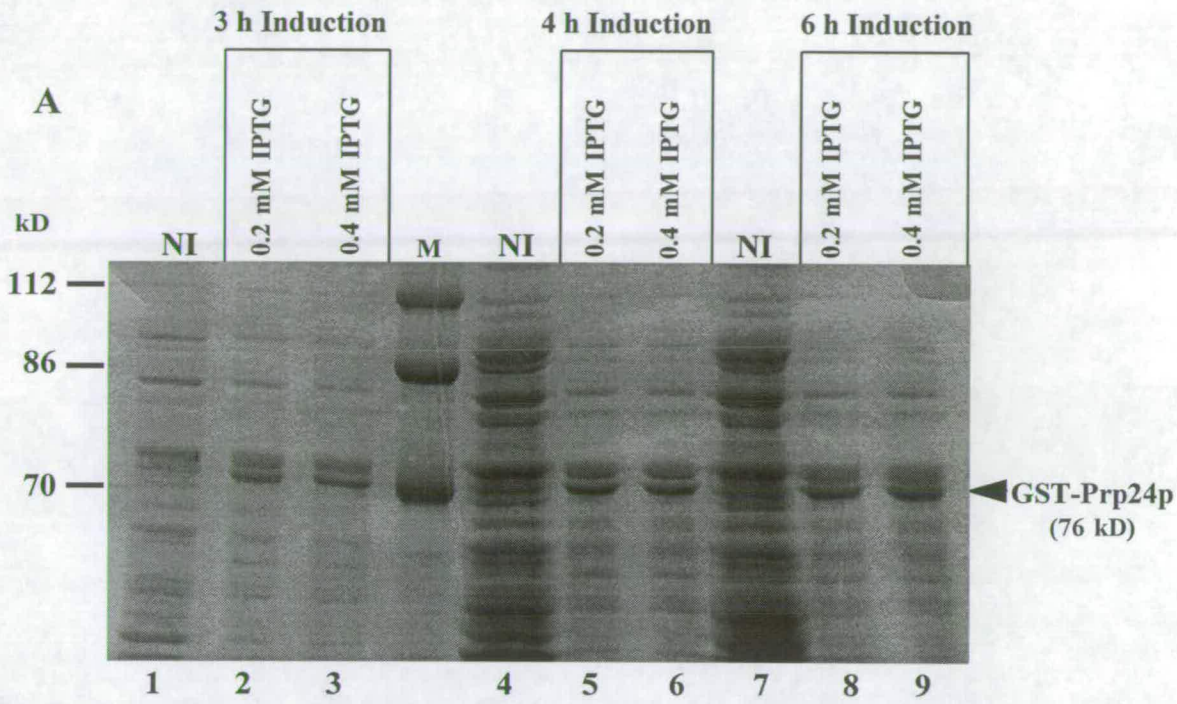
***Production of the glutathione-S-transferase-Prp24 fusion protein with pGEX24:*** *BMH71-18* (pGEX24) cells (Tables 6 and 9) were induced for the production of the GST-Prp24 fusion protein by addition of IPTG (section 2.9.1). After extraction and fractionation on an SDS-polyacrylamide gel, proteins were visualised by Coomassie blue staining (Figure 15, Panel A). A polypeptide of the expected size of the fusion protein (~76 kD) was produced only when IPTG was added in the culture. The best production of GST-Prp24 fusion protein was obtained after 4-6 h incubation in the presence of (0.2 mM-0.4 mM) IPTG (lanes 5, 6, 8, 9). Increasing the IPTG concentration (0.2 mM compared to 0.4 mM) in the culture had no effect on the production of the fusion protein which was limited. The induction of GST-Prp24p production reduced the growth rate of the cells decreasing overall protein production (compare lane 4 with 5, 6 and lane 7 with 8, 9), suggesting that GST-Prp24p was toxic in *E.coli*. This meant only small amounts of the fusion protein were produced.

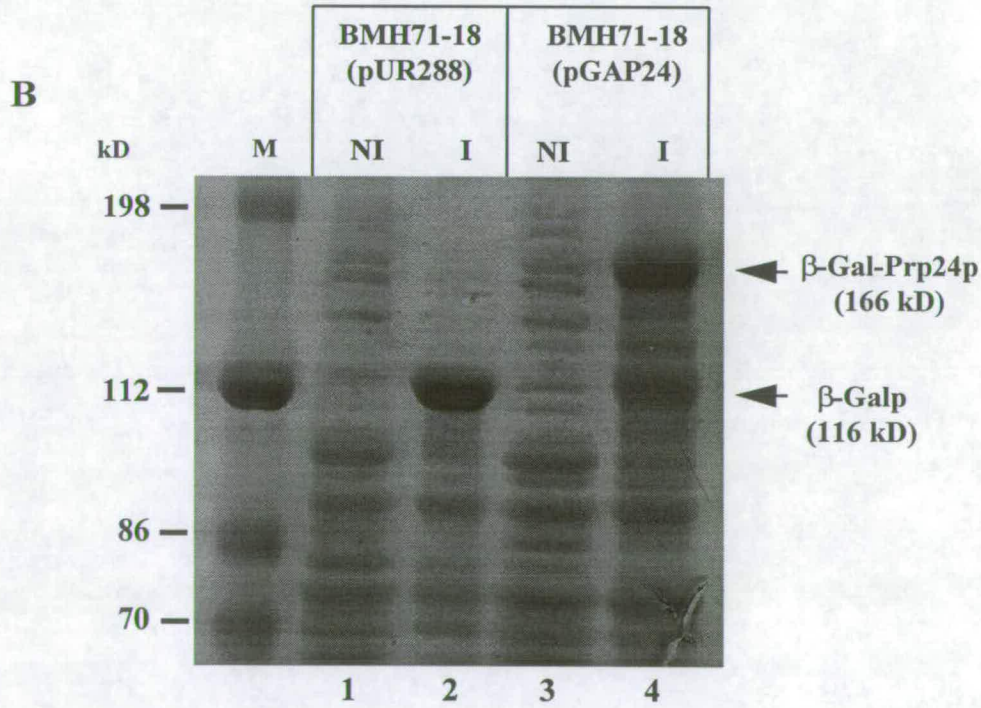
***Production of the β-galactosidase-Prp24 fusion protein with pGAP24:*** *BMH71-18* (pGAP24) cells were induced for the production of the β-galactosidase-Prp24 fusion protein by addition of IPTG (the control culture had no IPTG). After 4 h of induction cells were collected, proteins extracted, resolved on a 8% SDS-polyacrylamide gel and detected by Coomassie blue staining (Figure 15, Panel B). The fusion protein was produced only in the presence of IPTG and resolved at the expected size of 166 kD (compare lanes 3 and 4). A protein of a size similar to β-galactosidase was also detected, indicating that the fusion protein was partially unstable (lane 4). This partial degradation of the fusion protein has been mentioned by Rüther *et al.* (1983) but the reasons for this proteolysis are not well understood. The fusion protein was not toxic in bacteria and no major contaminating polypeptides were detected by Coomassie blue staining. A large quantity of proteins was extracted from induced cells (section 2.5.2) and the β-galactosidase-Prp24 fusion protein was partially purified by electroelution (section 2.9.2) and used to immunise rabbits against Prp24p.

**Figure 15. Analysis of proteins in E. coli cell extracts by Coomassie blue staining.**

**A:** Fresh cultures (5 ml) of *E. coli* BMH71-18 (pGEX24) cells were grown to $OD_{595}$=0.3-0.4 and the production of the GST-Prp24 fusion protein was induced by addition of 0.2 mM IPTG (lanes 2, 5, 8) or 0.4 mM IPTG (lanes 3, 6, 9). Control lacking IPTG were included (lanes 1, 4, 7). Cells were collected after 3 h, 4 h or 6 h induction. Proteins were extracted (small scale method, section 2.5.2) and a 20 µl aliquot (undetermined protein concentration) was loaded and fractionated on a 6% SDS-polyacrylamide gel. Proteins were detected by Coomassie blue staining. The expected size of the fusion protein was 76 kD. M: molecular weight markers. NI; noninduced

**B:** Competent BMH71-18 *E. coli* cells were transformed with pGAP24 (lanes 3, 4) or with pUR288 vector (lane 1, 2). Fresh culture (5 ml) was grown to $OD_{595}$=0.3-0.4, then IPTG was added to a final concentration of 1 mM IPTG (lanes 2 and 4) or was omitted (lanes 1 and 3). After 4 h induction, cells were collected and proteins extracted (section 2.5.2). 20 µl aliquot (undetermined protein concentration) was loaded on a 6% SDS-polyacrylamide gel. Proteins were detected by Coomassie blue staining. The expected size of the fusion protein was 166 kD and of the β-galactosidase was 116 kD. M: molecular weight markers. NI: no IPTG added; I: IPTG added.


A
3 h Induction
4 h Induction
6 h Induction
0.2 mM IPTG
0.4 mM IPTG
0.2 mM IPTG
0.4 mM IPTG
0.2 mM IPTG
0.4 mM IPTG
kD
NI
M
NI
NI
112
86
70
GST-Prp24p
(76 kD)
1 2 3 4 5 6 7 8 9


B
BMH71-18
(pUR288)
BMH71-18
(pGAP24)
kD
M
NI
I
NI
I
198
112
86
70
β-Gal-Prp24p
(166 kD)
β-Galp
(116 kD)
1 2 3 4

***Production of the 6X histidine-Prp24 fusion protein with pIP24:*** BMH71-18 (pIP24) cells were induced for the production of the 6X His-Prp24 fusion protein by addition of IPTG. After 4h induction (0.4 mM IPTG) cells were collected, proteins extracted and fractionated on an 8% SDS-polyacrylamide gel. The detection of the proteins by Coomassie blue staining indicated that the fusion protein was produced only in the presence of IPTG and possessed the expected size (50 kD) (data not shown). High yields of protein were produced and used to affinity purify the polyclonal anti-Prp24 antibodies (section 2.9.5).

### 3.3.2 Detection of Prp24 protein by immunoblotting

Anti-Prp24 antibodies were obtained from rabbits after six injections with the β-galactosidase-Prp24 fusion protein. Two independant batches (1015 and 1016) of antiserum were obtained from the immunization of two rabbits, both were tested for their ability to recognize Prp24p in a total yeast protein extract by Western blotting (section 2.5.6) (Figure 16, Panel A). In extract prepared from cells overproducing Prp24p, several polypeptides were detected at around 50 kD which may have been Prp24p (lane 2). Batch 1015 had more cross-reacting bands than 1016.

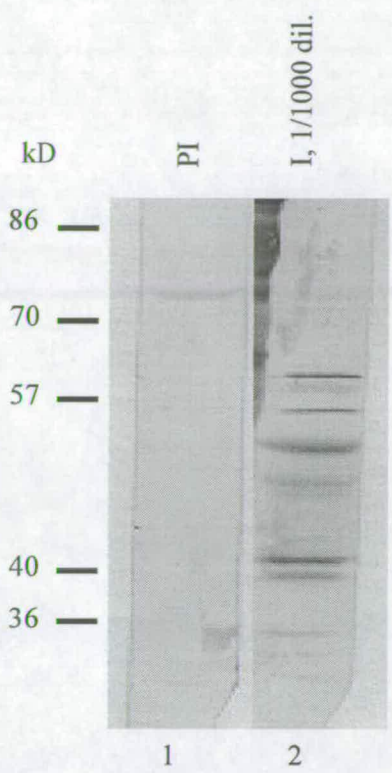
Proteins immunoprecipitated with anti-Prp24 antisera bound to PAS beads were detected by Western blotting (Figure 16; Panel B). Both batches of antibodies specifically precipitated Prp24p under nondenaturing conditions (section 2.8), however the protein was compressed by the heavy chain of the IgG (~60 kD) and consequently possessed an apparent molecular weight of ~45 kD (compare lane 1 with lanes 3 and 5). A protein a size similar to the major signal detected from a total yeast extract was immunoprecipitated with both batches of antibodies covalently attached to an Affi-Gel HZ resin (section 2.9.6) (data not shown). No difference in the efficiency of immunoprecipitation of Prp24p was noticed between the two antisera tested, also both immunoprecipitated Prp24p with similar efficiency after protein denaturation (data not shown).

**Figure 16. Detection by Western blotting of Prp24p with both batches of anti-Prp24 rabbits antisera.**

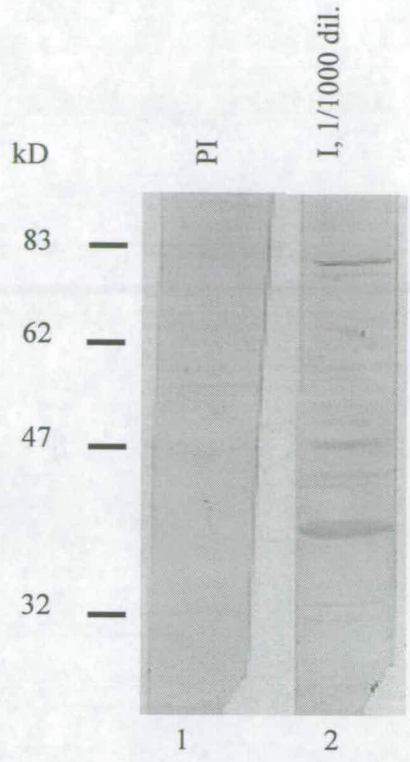
**A:** Yeast S150-2B (pBM24: *CEN4*, *URA3*, *GAL1-PRP24;* Tables 7 and 9) cells were grown in YMGlyLac medium (section 2.1.2.3) supplemented with tryptophan (Table 5) and galactose was added (20% (v/v) final) when cells reached the exponential phase. After 5 h galactose induction of *PRP24* expression, cells were harvested and proteins extracted (section 2.5.2). The protein content was determined by Bradford assay (section 2.5.1) and 100 µg of total proteins were loaded in each slot of a 8% SDS-polyacrylamide gel. Proteins were transferred to a PVDF membrane by Western blotting (section 2.5.6), detected by Ponceau S staining and the membrane was cut into several bands each containing the same amount of protein. A band was incubated with the preimmune serum (1/1000 fold diluted in 5% BSA/1X PBS) (lane 1) or with the crude anti-Prp24 antiserum (1/1000 fold diluted in 5% BSA/1X PBS) (lane 2).

**B:** Detection by Western blotting of immunoprecipitated proteins with anti-Prp24 antibodies. A splicing extract (60 µl; 4 mg total proteins) prepared from S150-2B (pBM24) cells collected after 5 h galactose induction of Prp24p production was mixed with an equal volume of 2X IQ buffer (section 2.8). The mixture was incubated with crude antibodies attached to PAS beads (final bleed, lanes 3, 5) or with preimmune serum (lanes 2, 4). Immunoprecipitation was carried out for 2h at 4°C (section 2.8), after washes with buffer containing 150 mM NaCl, immunoprecipitated proteins were recovered and loaded on a 8% SDS-polyacrylamide gel. Alongside, 100 µg of total proteins (provided from the same protein extract) was loaded on the gel (lane 1), proteins were transferred to a PVDF membrane and detected by Western blotting with anti-Prp24 serum batch 1016 (final bleed) 1/1000 diluted in 5% BSA/1X PBS. IP: immunoprecipitation.

A

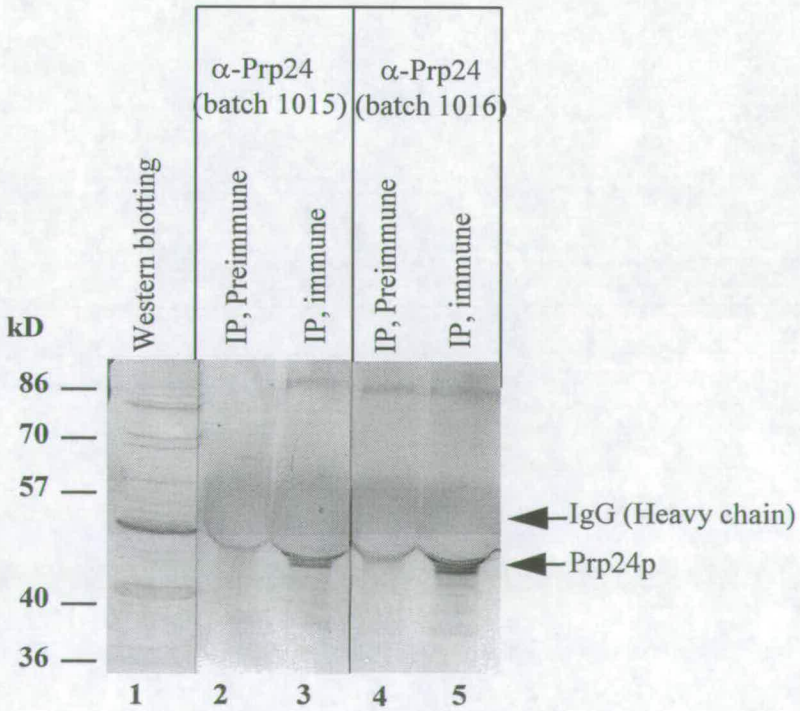



**Western blotting with α-Prp24 Abs (batch 1015)**




**Western blotting with α-Prp24 Abs (batch 1016)**

B




**Immunoprecipitation with α-Prp24 Abs**

### 3.3.3 Analysis of snRNAs co-immunoprecipitated with anti-Prp24 antiserum

The snRNAs that are co-immunoprecipitated with anti-Prp24 antisera from a yeast splicing extract (SC261, Table 7) are shown in Figure 17. Both batches of antisera precipitated mainly U6 snRNA and only a very tiny fraction of U4 snRNA, while no U1, U2 or U5 snRNAs were detected. Neither the incubation at 25°C for 25 min nor the induction of spliceosome formation affected the immunoprecipitation of snRNAs by either anti-Prp24 antiserum, demonstrating that both batches of antibodies detected Prp24p associated with the U6 snRNA before spliceosome formation. This was in agreement with the data of Shannon and Guthrie (1991) and Jandrositz and Guthrie (1995), showing that the immunoprecipitated Prp24p was only associated with the U6 snRNA in the U6 snRNP particles (in a wild-type yeast background).

## 3.4 Discussion

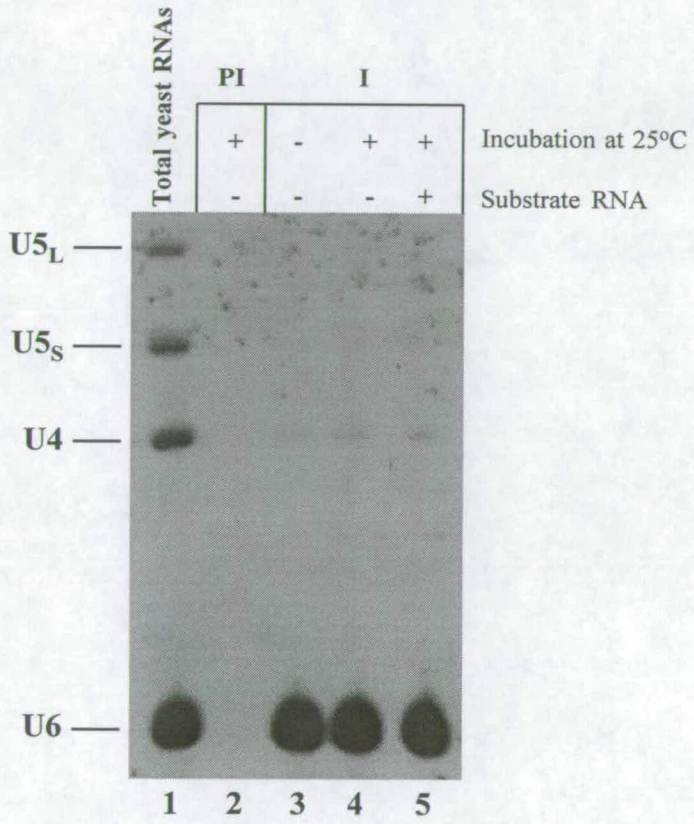
It was shown that Prp24p was essential for yeast cell viability, in agreement with data reported subsequently by Vaidya *et al.* (1996).

The production in *E. coli* of a β-galactosidase-Prp24 fusion protein allowed the partial purification of sufficient amounts of protein to induce an immune response in rabbits. Two batches of polyclonal anti-Prp24 antisera were produced. Both antibodies recognized a polypeptide of an apparent molecular weigth of 50 kD in a yeast extract overproducing Prp24p. Much less cross-reacting proteins were detected with antiserum batch 1016. Both anti-Prp24 antibodies immunoprecipitated with an equal efficiency Prp24p (~50 kD) under nondenaturing conditions (150 mM NaCl) or after protein denaturation.

In our hands, anti-Prp24 antibodies co-immunoprecipitated mainly the U6 snRNA and traces of the U4 snRNA. This was demonstrated in a splicing extract prepared from wild-type yeast cells in a reaction containing ATP. Spliceosome formation was not required for Prp24p-U6 snRNA association. The same results were obtained when

**Figure 17. Northern blotting analysis of snRNAs co-immunoprecipitated with anti-Prp24 antisera.**

A splicing extract (40 µl) prepared from the wild-type yeast SC261 strain (Table 7) was present in a splicing reaction mix (100 µl), and incubated for 25 min at 25°C (or left on ice) (section 2.6.2.1). Spliceosome formation was induced by addition to the sample of substrate RNA (1 µM of actin p283, Table 9). No spliceosomes were formed when substrate RNA was omitted. After 25 min incubation, samples where quenched on ice and Prp24p was immunoprecipitated under nondenaturing conditions (section 2.8) with anti-Prp24 antisera batches 1016 (final bleed) or with the preimmune serum. After five washes containing 150 mM NaCl, co-immunoprecipitated snRNAs were recovered (section 2.8) and transferred to a nylon membrane (section 2.4.4.3). The presence of U1, U2, U4, U5 and U6 snRNAs was detected by hybridization with end-radiolabelled (section 2.4.3.10) oligodeoxynucleotides 486A, G1803, 483A, 485A and Taq6A respectively (Table 8). Total yeast RNA was extracted from a 5 µl splicing extract (lane 1) (section 2.8). The membrane to which U1 snRNA and U2 snRNAs were transferred is not shown but no immunoprecipitation was detected. Identical results (not shown on this Figure) were obtained when anti-Prp24 antibodies batch 1015 (final bleed) was used.




**SnRNAs Co-IP with anti-Prp24**
**Abs batch 1016**

affinity purified anti-Prp24 antibodies were used (see section 5.2.1). These results suggest that the polyclonal antibodies recognize Prp24p interacting with the U6 snRNA in the U6 snRNP particle. This is in agreement with data of Shannon and Guthrie (1991) and Jandrositz and Guthrie (1995). These authors demonstrated that Prp24p was associated with the U6 snRNA in the U6 snRNP particle only. Recently Prp24p has been proposed to be required for the recycling of the U4/U6 snRNP after spliceosome disassembly, in an ATP-dependent manner (Raghunathan and Guthrie, 1998). These authors showed that in yeast cells producing a Prp24-3HA fusion protein, U6 snRNA and also U4 snRNA were associated with Prp24-3HA in an ATP-dependent manner (Raghunathan and Guthrie, 1998). Several explanations for the discrepancy between these published data and this work are proposed:

1- Raghunathan and Guthrie (1998) have investigated the association of Prp24p with snRNAs at low salt concentrations (50 mM NaCl), while in this work more stringent conditions (150 mM NaCl). Therefore in the presence of 150 mM NaCl, the association between Prp24p and the U4/U6 duplex may have been disrupted.

2- Although these authors demonstrated that the Prp24-3HA protein was functional, the addition of extra amino acids might have increased the affinity of Prp24p for the U4/U6 duplex.

3- The association of Prp24p with the U4/U6 snRNAs might have been disrupted by the binding of the polyclonal antibodies used, or the Prp24p epitope(s) was masked in the Prp24p-U4/U6 snRNP complex.

The ability of both antisera to immunoprecipitate Prp24p UV-crosslinked to the radiolabelled U6 RNA has been investigated and the results are presented in Chapter 5 (Figure 22).

# CHAPTER FOUR

## The U6 snRNP RECONSTITUTION

# INTRODUCTION

SnRNPs have been reconstituted *in vitro* in both mammalian and yeast systems allowing the study of protein-protein, protein-RNA and RNA-RNA interactions occuring in snRNP particles and in the spliceosome (for examples see McPheeters *et al*, 1989; Wersig *et al.*, 1992; Ségault *et al.*, 1995; O'Keefe *et al.*, 1996).

The U6 snRNP reconstitution method (section 2.6.2.3) has been published by Fabrizio *et al.* in 1989. The principle of this technique consists of replacing, in a yeast splicing extract, the endogenous U6 snRNA by one produced by transcription *in vitro*. For this purpose a U6 knockout oligo (Table 8) complementary to the U6 snRNA sequence $U_{28}$-$U_{54}$ was added to the yeast cellular extract and allowed to anneal to the endogenous U6 snRNA. The region of the U6 snRNA involved in the RNA-DNA interaction was consequently degraded by the RNase H which is present in the yeast cellular extract and is activated upon incubation at 30°C. Due to the depletion of the endogenous U6 snRNA, the splicing activity of the extract was abolished. The addition of U6 snRNA produced by transcription *in vitro* reconstitutes functional U6 snRNP particles and therefore restores the splicing activity of the yeast extract.

Although this reconstitution technique appears to be relatively simple, numerous difficulties have been encounted which will be described and discussed in this chapter.

## 4.1 Preparation of a Heat Stable Yeast Splicing Extract and Test for Depletion of the Endogenous U6 SnRNA

To deplete the U6 snRNA from the yeast cellular extract, it was necessary to incubate the reconstitution mixture at 30°C for 30 min in order to activate the endogenous RNase H (section 2.6.2.3). Comparison of the splicing activity of both heat treated and untreated yeast extracts, in the absence of U6 knockout oligo, revealed that

the incubation at 30°C was abolishing completely or partially the splicing activity of the extract. The splicing conditions reported by Fabrizio *et al.* (1989) were different from those described in section 2.6.2.1. Whereas Fabrizio *et al.* used a potassium phosphate buffer of pH 7.0 and 2 mM spermidine in their splicing reaction, potassium phosphate buffer pH 7.5 was used and no spermidine was added in the reaction. These differences did not affect the heat stability of the extract, no improvement was obtained. The splicing extracts were prepared from the wild-type yeast strain S150-2B (Table 7), extracts from a yeast strain containing mutations in protease encoding genes (SC261: *prb*1-1132, *prc*1-407, *pep*4-3; Table 7) were prepared; also the effect of the addition of protease inhibitors during the extract preparation was tested, no changes were noticed. The exclusive use of sterile solutions and material (centrifugation tubes, Pasteur pipettes, magnetic stirrers...) and the careful resuspension of yeast spheroplasts (spheroplasts were first resuspended in a small volume of the appropriate buffer until complete homogenisation, then more buffer was added), were the two determining precautions which permitted the preparation of heat stable splicing extracts.

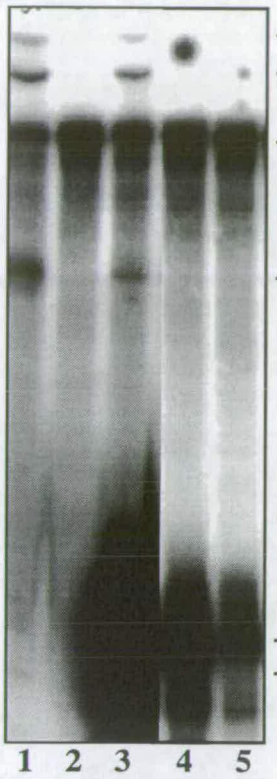
An accurate titration of the knockout oligo was essential to ensure complete digestion of the endogenous U6 snRNA and to avoid degradation of the transcribed U6 RNA by excess oligo. For most of the extracts tested, a final concentration of 110 nM-140 nM of U6 knockout oligo was sufficient to completely abolish the splicing activity of the extract. However when another knockout oligo was included to deplete either U2 or U5 snRNAs in addition to the U6 snRNA (sections 2.6.2.4 and 2.6.2.5), the stability of the added U6 RNA was compromised. Consequently the efficiency of the U6 snRNP reconstitution was reduced. The splicing activity of a yeast extract before and after U6 snRNP reconstitution is shown in Figure 18 (Panel A). In addition to the radiolabelled splicing intermediates and products, the U6 RNA ($^{32}$P-αATP) was also detected. The splicing activity recovered after U6 snRNP reconstitution varies from ~ 50% of the initial level of activity of the extract (based on the comparison of the lariat-intron-exon 2 and lariat-intron signals), or was ~20% of the initial splicing activity (based on the comparison of the spliced RNA signal) (compare lanes 1 and 3). In this particular reconstituted yeast cellular extract, the spliced RNA was less stable than the lariat-

**Figure 18. Analysis of the U6 RNA transcribed *in vitro* on denaturing gel.**

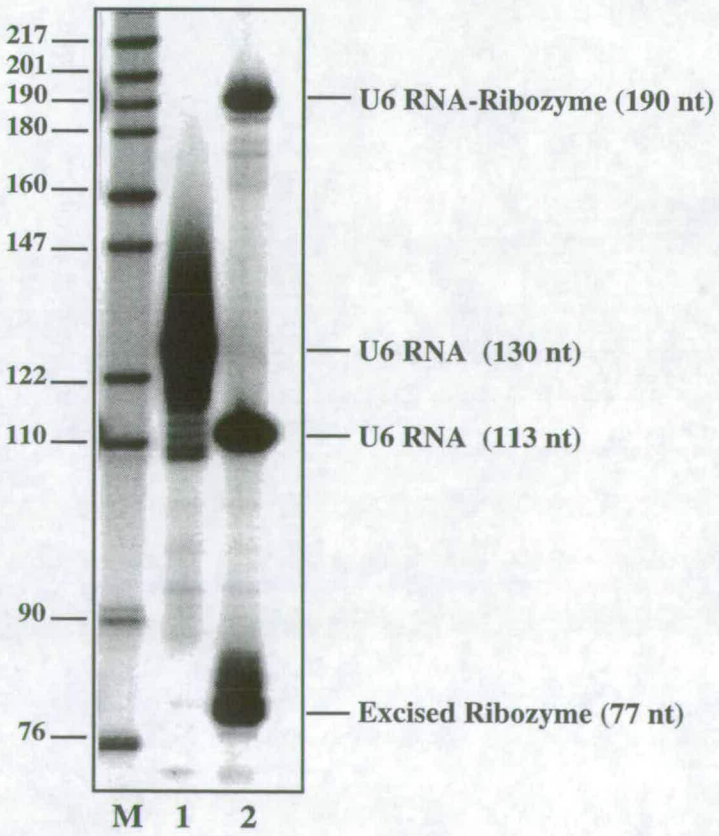
**A:** U6 snRNP reconstitution and splicing reaction. The initial splicing activity of the yeast cellular extract is shown in lane 1, the extract has been preincubated at 30°C for 30 min prior to the splicing reaction. The addition of 140 nM of U6 knockout DNA oligo was sufficient to completely deplete the endogenous U6 snRNA and abolish the splicing activity of the extract (lane 2). U6 snRNPs were reconstituted with a radiolabelled U6 snRNA produced by transcription *in vitro*. In each reconstitution reaction 6 nM of U6 RNA was added (lanes 3-5). The recovery of the splicing activity was observed in lane 3. In addition to the endogenous U6 snRNA, U2 snRNA was depleted by the presence of 600 nM of U2 knockout oligo SRU2 (Table 8) (lane 4). The U5 snRNA was also degraded in addition to the U6 snRNA with 700 nM of each U5 knockout oligos Rok8L and Rok8R (Table 8) (lane 5). A 10 µl reconstitution reaction was carried out with a splicing extract prepared from the yeast strain SC261 (Table 7; lanes 1-4) or from the yeast strain Rok8 (provided by I. Dix, this lab) (Table 7; lane 5). After splicing reaction, RNAs were recovered and resolved on a 7% polyacrylamide/urea/TBE gel. RNA signals were detected by autoradiography. IVS-E2: lariat-intron-exon 2; IVS: lariat-excised intron; pre-mRNA: actin substrate RNA (p283, Table 9); mRNA: spliced RNA; E1: exon 1; U6 RNA: U6 snRNA produced by transcription *in vitro* with the T7 RNA polymerase.

**B:** U6 RNA was transcribed (10 µl reaction) *in vitro* from the template pUC12U6 linearized with *Dra*I (section 2.4.4.1) (lane 1) or from the pU6ribo template linearized with *Eco*RI (lane 2). Transcripts were resolved on a 6% polyacrylamide/urea/TBE gel and detected by autoradiography. A DNA ladder (pBR322 plasmid digested with *Msp*I and radiolabelled with $\alpha$-$^{32}$PdCTP) was run alongside with the U6 snRNA transcripts. U6 RNA: U6 snRNA produced by transcription *in vitro* with the T7 RNA polymerase and radiolabelled with $^{32}$P-$\alpha$ATP. M: DNA ladder.

A


IVS-E2
IVS
pre-mRNA
mRNA
U6 RNA
E1
1 2 3 4 5

B


217
201
190
180
160
147
122
110
90
76
U6 RNA-Ribozyme (190 nt)
U6 RNA (130 nt)
U6 RNA (113 nt)
Excised Ribozyme (77 nt)
M 1 2

intron product. As expected no splicing activity was detected when the endogenous U2 or U5 snRNA was depleted in addition to the U6 snRNA (lanes 4 and 5), however under these experimental conditions much less radiolabelled U6 RNA was recovered (compare the U6 RNA signal from lane 3 with lanes 4 and 5). The U6 RNA was apparently degraded in the extract. This U6 RNA instability was also noticed when an oligodeoxynucleotide, lacking any complementarity to any of the known snRNAs, was added to the reconstitution reaction with the U6 knockout oligo (data not shown). This suggested that the increased degradation of the U6 RNA was not dependent on the sequence of the second oligo added.

Fabrizio *et al.* demonstrated that after 30 min incubation at 30°C less than 1% of the initially added U6 knockout oligo remained in the extract, having probably been degraded by endogenous DNase activity. Therefore presumably, in the double knockout experiment the presence of a second oligo limited the DNase degradation of the U6 knockout oligo. Consequently a larger fraction of the U6 knockout oligo remained in the extract and annealed to the U6 RNA which was then RNase H digested. Therefore when the endogenous U2 or U5 snRNAs were depleted in addition to the U6 snRNA, a lower quantity of U6 knockout oligo was added in the reconstitution reaction to compensate for the longer life span of the U6 knockout oligo in the extract whereas the incubation time was unchanged.

## 4.2 Analysis of the U6 SnRNA Produced by Transcription *In Vitro*

The length of the U6 RNA transcribed from pUC12U6 or from pU6ribo (Table 9) was analysed on a denaturing gel (Figure 18, Panel B). The U6 RNA transcribed from the pU6ribo template was 190 nucleotides long. It possesses at its 3' end a hairpin ribozyme which autocleaves and releases the ribozyme (77 nt) and the U6 RNA of the expected size (113 nt) (lane 2). The U6 RNA transcribed from pUC12U6 was longer than expected. According to the DNA ladder, it was 130 nt in length and therefore possesses 17 extra nucleotides (lane 1).

The unexpected production of T7 transcripts containing extra nucleotides at their 3' end has been reported in the literature (Cazenave *et al.*, 1994; Triana-Alonso *et al.*, 1995). Triana-Alonso *et al.* suggested a reduction in the UTP concentration to circumvent this extension problem. A decrease in the concentration of UTP had no effect on the size of the U6 RNA transcribed. Likewise, various magnesium concentrations were tested, in combination with different nucleotide concentrations; the U6 RNA produced by transcription *in vitro* from pUC12U6 was still 130 nt long. In an attempt to disrupt any U6 RNA secondary structures which may have been responsible for this 3' end extension, DMSO was added in the transcription reaction at 5% (v/v) or 10% (v/v). No improvement was observed. No changes were detected when the reaction temperature was changed to 30°C or 40°C, or when different sources of T7 RNA polymerases were tested (Boehringer Mannheim, New England Biolabs, Pharmacia). A PCR product containing the U6 snRNA gene downstream of the T7 transcription promoter was produced using oligodeoxynucleotides S5760 and S5761 (Table 8). When this was used as a DNA template in the transcription reaction, the resulting U6 RNA was again 130 nt in length. Whatever the experimental conditions tested, the U6 RNA produced by transcription *in vitro* was always longer than expected.

Cazenave and Uhlenbeck (1994) have shown that T7 RNA polymerase is capable of extending the 3' end of a transcript by RNA template-directed RNA synthesis. This occurs because the RNA made from the DNA template is able to form an intra or intermolecular primed-template. The same 3' end extention was indeed occuring when the U6 RNA was transcribed *in vitro*. I. Kelly (University of Cambridge) sequenced the 17 extra nucleotides located at the 3' end of the U6 RNA. The sequence was 5'-UAAAAAAACGGUUCAUC-3' (I. Kelly, personal communication), the underlined sequence is complementary to U6 snRNA nucleotides 78-90 confirming this RNA-templated transcription activity.

## 4.3 U6 SnRNP Reconstitution with the U6 RNA Transcribed *In Vitro*

The fact that the U6 RNA was longer than expected was a concern for its functionality and the success of the U6 snRNP reconstitution. The ability of the 130 nt-U6 RNA to reconstitute snRNP particles was investigated and compared with that obtained when the 113 nt-U6 RNA was used (Figure 19). The splicing extract tested was heat stable, and the heat treatment seems to increase slightly the stability of the spliced RNA and the lariat-intron signals (compare lanes 1 and 2). The amount (130 nM) of U6 knockout oligo added was sufficient to abolish the splicing activity of the extract (lane 3). After U6 snRNP reconstitution with the longer U6 RNA (130 nt), the initial splicing activity of the extract was recovered with an efficiency which varies from 60% (with 5 nM of U6 RNA, lane 4) to 100% (with 100 nM of U6 RNA, lane 7), with respect to the lariat-intron and lariat-intron-exon 2 signals. However slightly less than 100% of mRNA was recovered after U6 snRNP reconstitution with 50 nM or 100 nM of U6 RNA (130 nt) (compare mRNA signal in lane 2 with lanes 6 and 7). For no obvious reason the spliced RNA was instable. When U6 snRNPs were reconstituted with the U6 RNA of 113 nt in length, a slightly lower if not similar reconstitution efficiency was detected. Therefore the extra 17 nucleotides present at the 3' end of the 130 nt-U6 RNA do not affect its ability to reconstitute functional U6 snRNP particles.

In order to detect the U6 RNA after reconstitution and splicing reaction, U6 snRNP particles were reconstituted with a radiolabelled 130 nt-U6 RNA or 113 nt-U6 RNA (Figure 20, Panel A). An RNA, of a size similar to the 113 nt-U6 RNA initially added in the reaction, was detected after U6 snRNP reconstitution and splicing (compare lanes 1 and 2). When the 130 nt-U6 RNA was used, a shorter RNA (~113 nt in length) was detected, which was not initially added in the reaction (compare lanes 3 and 4). This shortened U6 RNA was also detected in the absence of substrate RNA, suggesting that no spliceosome formation was required for the 130 nt-U6 RNA to be shortened (data not shown). This result was unexpected. Was the 130 nt-U6 RNA (or at least a fraction of it) processed in the extract to an RNA of the expected size? Were both U6 RNA species functional?

**Figure 19. The 130 nt-U6 RNA is functional.**

U6 snRNPs were reconstituted with the 130 nt-U6 RNA (lanes 4-7) or with the 113 nt-U6 RNA (lanes 8-11). The splicing activity of the extract prepared from SC261 (Table 7) is shown in lane 1. Its ability to splice was not altered when the extract was preincubated at 30°C for 30 min before splicing reaction (lane 2). The splicing activity of the yeast extract was abolished when 130 nM of U6 knockout oligo was added to deplete the endogenous U6 snRNA (lane 3). Increasing amounts of U6 RNA were added to reconstitute U6 snRNPs. IVS-E2: lariat-intron-exon 2; IVS: lariat-excised intron; pre-mRNA: actin substrate RNA (p283, Table 9); mRNA: spliced RNA; E1: exon 1.

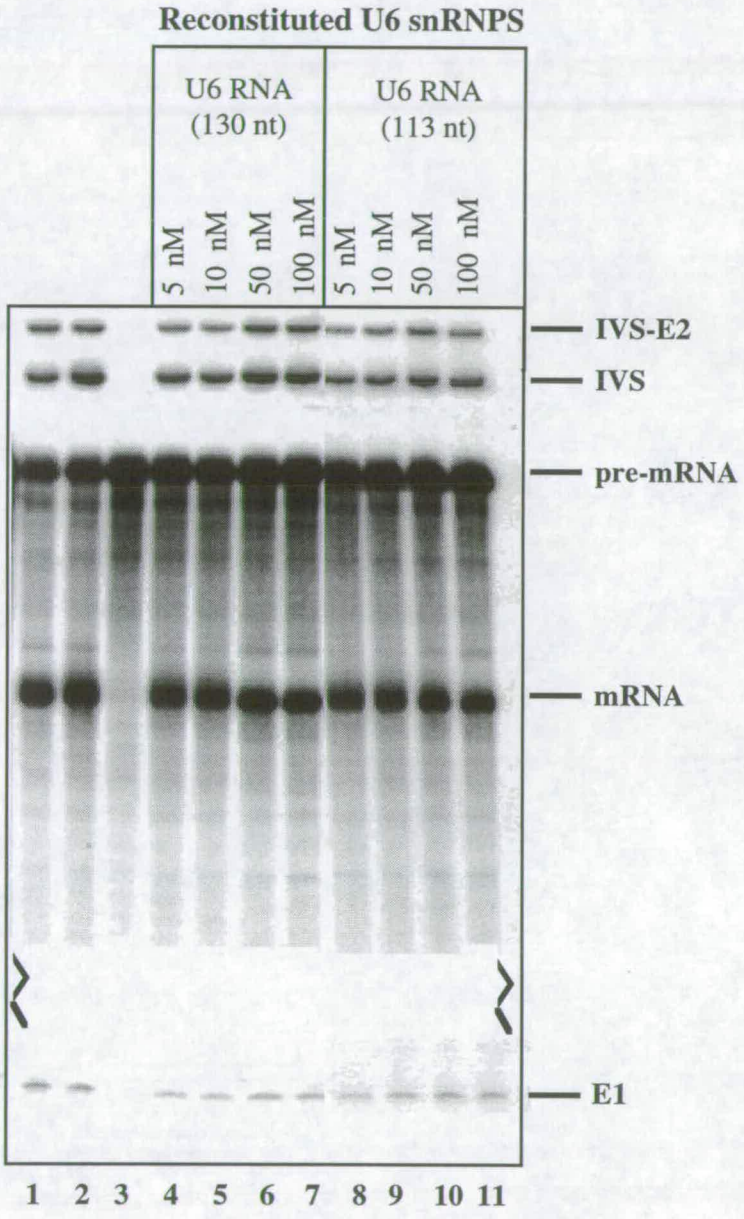
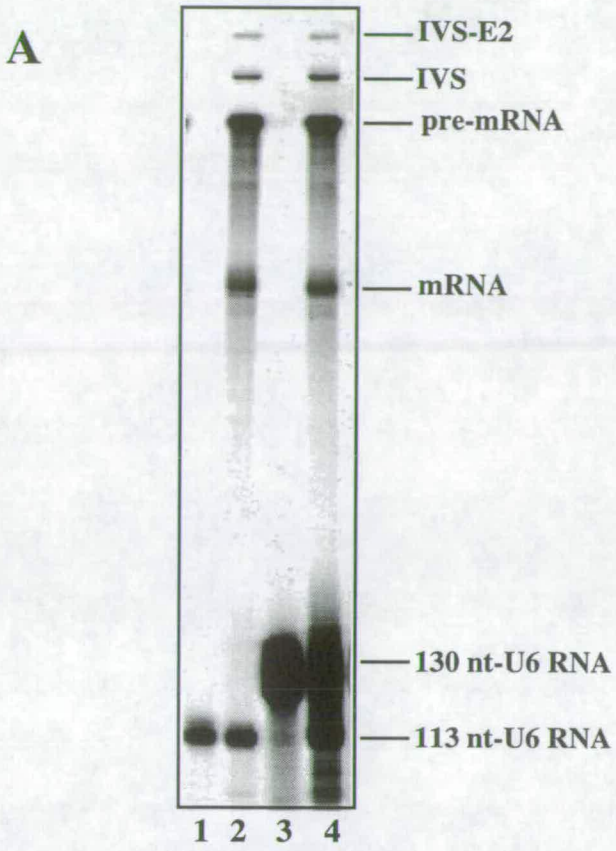
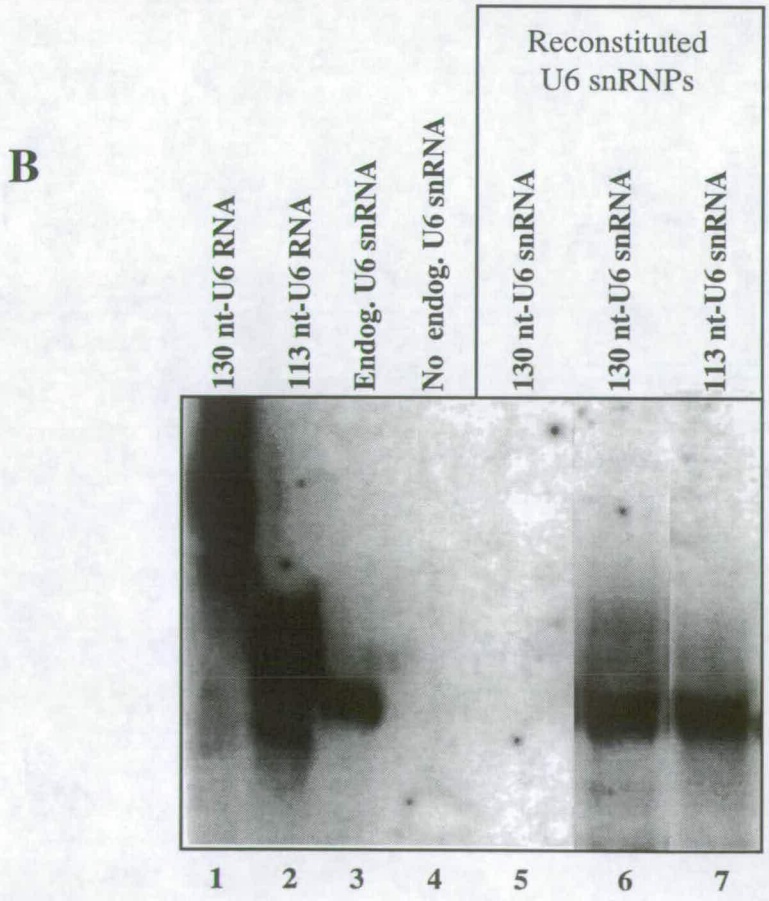

Reconstituted U6 snRNPS
U6 RNA (130 nt)
U6 RNA (113 nt)
5 nM
10 nM
50 nM
100 nM
5 nM
10 nM
50 nM
100 nM
IVS-E2
IVS
pre-mRNA
mRNA
E1
1 2 3 4 5 6 7 8 9 10 11

Figure 20. Analysis of the 130 nt-U6 RNA in a splicing extract.

**A:** Splicing activity of an extract prepared from SC261 cells (Table 7) assayed after U6 snRNP reconstitution. U6 snRNPs were reconstituted with a radiolabelled 113 nt-U6 RNA (30 nM) (lane 2) or with a radiolabelled 130 nt-U6 RNA (100 nM) (lane 4). 10 µl of reconstitution and splicing reactions were performed. After splicing, RNAs were recovered and resolved on a 7% polyacrylamide/urea/TBE gel. On the same gel 300 fmole of the 113 nt-U6 RNA (lane 1) and 1 pmole of the 130 nt-U6 RNA were loaded alongside. IVS-E2: lariat-intron-exon 2; IVS: lariat-excised intron; pre-mRNA: actin substrate RNA (p283, Table 9); mRNA: spliced RNA; E1: exon 1; U6 RNA: U6 snRNA produced by transcription *in vitro* with T7 RNA polymerase.

**B:** Analysis of the U6 RNA in the reconstituted Uss1p-containing snRNPs. U6 snRNP were reconstituted (50 µl reaction) with 80 nM of 130 nt-U6 RNA or 113 nt-U6 RNA. After reconstitution the U6 snRNA contained in U6 snRNP, U4/U6 snRNP and in U4/U6.U5 tri-snRNP particles was co-immunoprecipitated with the crude anti-Uss1 antisera under native conditions (lanes 6 and 7). As a control reconstituted snRNP particles were incubated with PAS beads only (lane 5). Immunoprecipitation and washes were carried out in 150 mM NaCl. After immunoprecipitation the RNAs were recovered, fractionated on a 6% polyacrylamide/urea/TBE gel, blotted (section 2.8) and the U6 snRNA was probed with a radiolabelled oligodeoxynucleotide Taq6A (Table 8; sections 2.4.3.10 and 2.4.3.11). Total RNAs were also recovered from 5 µl of splicing extract (lane 3) or from 5 µl of splicing extract lacking the endogenous U6 snRNA (lane 4). On the same gel 50 fmole of the initial stock of 130 nt-U6 RNA (lane 1) and 113 nt-U6 RNA (lane 2) were loaded alongside the other samples.


A
IVS-E2
IVS
pre-mRNA
mRNA
130 nt-U6 RNA
113 nt-U6 RNA
1 2 3 4


B
Reconstituted
U6 snRNPs
130 nt-U6 RNA
113 nt-U6 RNA
Endog. U6 snRNA
No endog. U6 snRNA
130 nt-U6 snRNA
130 nt-U6 snRNA
113 nt-U6 snRNA
1 2 3 4 5 6 7

The length of the U6 RNA in reconstituted snRNP particles was investigated. For this purpose U6 snRNP particles were reconstituted with the 130 nt-U6 RNA or with the 113 nt-U6 RNA. Then, U6 snRNP, the U4/U6 snRNP and to a lesser extent the U4/U6.U5 tri-snRNP were immunoprecipitated with the crude anti-Uss1 antiserum and the co-immunoprecipitated U6 RNAs were recovered and analysed by Northern blotting (section 2.4.4.3) (Figure 20, Panel B). The RNase H degradation of the endogenous U6 snRNA was total (lane 4), and the co-immunoprecipitation of the U6 RNA was antiserum specific (lane 5). Little size difference was observed between the transcribed 113 nt-U6 RNA used to reconstitute U6 snRNPs and the U6 RNA recovered from the Uss1p-containing snRNP particles (compare lanes 2 and 7). After U6 snRNP reconstitution with the 130 nt-U6 RNA, only a shorter (~113 nt) U6 RNA was recovered in Uss1p-containing snRNP particles (compare lane 1 and 6). This suggests that only the shortened 113 nt-U6 RNA was present in the yeast Uss1p-containing snRNPs. This could explain why no difference in the reconstitution efficiency was observed between the 113 nt-U6 RNA and the 130 nt-U6 RNA (see Figure 19).

## 4.4 Production of Large Amounts of Radiolabelled U6 SnRNA by Transcription *In Vitro*

Large amounts of highly radiolabelled U6 RNA were required to investigate U6 RNA-protein interactions (section 2.6.2.3). Initially the transcription recipe described in section 2.4.4.1 using $^{32}$P-αUTP to radiolabel the transcript was followed, and the plasmid pUC12U6 (Table 9) linearised with *Dra*I was used as template. The transcription efficiency obtained was very low while much larger quantities of U5 RNA were produced following the same recipe (data not shown). Thus the limitation of the transcription efficiency was specific for the U6 RNA. Several parameters (temperature, source of T7 RNA polymerase, addition of 5% (v/v) or 10% (v/v) DMSO, different concentrations of ribonucleotides, different transcription buffers) were modified in order to increase the production of radiolabelled U6 RNA; very little effect was

observed. The best transcription efficiency was obtained when U6 RNA was radiolabelled with $^{32}$P-αATP instead of $^{32}$P-αUTP (data not shown). Indeed labelling the U6 RNA with $^{32}$P-αUTP limits the availability of the UTP in the reaction; since U6 RNA contains a U at position 2, this may affect its transcription efficiency by causing high levels of abortive transcription. The first A in U6 RNA occurs at position 8, therefore limitating ATP in the reaction was probably less detrimental to the transcription efficiency of the full length U6 RNA.

A U6 RNA transcription template harboring at its 5' end six extra nucleotides (5'-GGGAGA-3') located between the T7 RNA polymerase consensus promoter sequence and the U6 snRNA encoding sequence was generated by PCR (section 2.4.3.8) (L. Verdone, this lab). Oligodeoxyribonucleotides Lory1 and S5761 were employed as primers (Table 8) and the plasmid pUC12U6 was used as template DNA in the PCR reaction. The amplified DNA product was named PCU6. The addition of these extra nucleotides at the 5' end of the transcription template greatly improved the transcription efficiency which was much higher than that obtained when the pUC12U6 was used as template DNA (L. Verdone, personnal communication). Large amounts of U6 RNA transcribed from the PCU6 template and radiolabelled with either $^{32}$P-αATP, $^{32}$P-αCTP or $^{32}$P-αUTP were also obtained, whereas only a U6 RNA radiolabelled with $^{32}$P-αATP was produced in reasonable amounts from the pUC12U6 template (data not shown). Further analysis indicated that the U6 RNA harboring six extra nucleotides at its 5' end was able to reconstitute functional U6 snRNP particles (i.e. tested for their ability to restore splicing activity of an extract lacking the endogenous U6 snRNA) (L. Verdone, personnal communication).

## 4.5 Discussion

In our hands the U6 RNA produced by transcription *in vitro* from pUC12U6 linearized with *Dra*I was 130 nucleotides long instead of the expected 113 nucleotides. Seventeen extra nucleotides were added at the 3' end of the transcript, probably due to an RNA-template directed synthesis. Fabrizio *et al.* (1989) made no mention of this 3' end extension of the full length U6 transcript. The plasmid (pUC12U6) used for the run-off transcription was similar to the one used by Fabrizio *et al.*, (1989): both plasmids contain the T7 promoter followed by an extra G, and they both were linearized with *Dra*I the transcription of the full length U6 snRNA molecule. The only difference was the vector in which the T7-U6 RNA gene was cloned: pUC12U6 plasmid was based on the vector pUC12, while Fabrizio *et al.* used the pUC18 vector. This difference was unlikely to be responsible for the transcription discrepancy. Also, the transcription conditions used by Fabrizio *et al.* (which were slightly different from those described in section 2.4.4.1) produced a 130 nt-RNA in our hands. This disparity remains obscure.

Fabrizio *et al.* (1989) showed that a longer U6 RNA containing 8 extra nucleotides at the 3' end, produced by run-off transcription of their plasmid DNA template linearized with *Bam*HI, gave a lower recovery of splicing activity. The functionality of the U6 RNA seems more affected by the sequence of the added nucleotides at the 3' end rather than by their number. Indeed, the results reported in this chapter showed that the addition of 17 nucleotides at the 3' end of the U6 RNA did not affect its functionality; the 130 nt-U6 RNA restored splicing activity with the same efficiency than the 113 nt-U6 RNA.

Another U6 transcription template has been generated by PCR which contains six extra nucleotides at its 5' end. This DNA template was named PCU6 and it allowed the production of large amounts of U6 RNA radiolabelled with either $^{32}$P-αATP, $^{32}$P-αCTP or $^{32}$P-αUTP. The presence of the 5'-GGGAGA-3' sequence at the 5' end of U6 RNA did not affect its functionality. This U6 RNA was therefore utilised to investigate U6

RNA-protein interactions (chapters 5 and 6). In the UV-crosslinking experiments presented in the following chapters, for convenience we will name the U6 RNA transcribed from pUC12U6/*Dra*I and radiolabelled with $^{32}$P-αATP: U6 RNA ($^{32}$P-αATP) and the U6 RNA transcribed from PCU6 and radiolabelled with $^{32}$P-αUTP: U6 RNA ($^{32}$P-αUTP).

The analysis of the splicing intermediates and products detected after U6 snRNP reconstitution and splicing reaction indicated that sometimes these RNA species were unstable. Optimization experiments of *in vitro* splicing showed that different pre-mRNAs were spliced at different rates, and also the products and intermediates were degraded at different rates depending partly on the particular splicing extract (M. Lossky, 1988). In our hands, for the same splicing extract tested, the stability of the splicing signals varies from one reaction sample to another. This instability was not well understood and may occur randomly during the reaction and/or during the recovery of the RNAs after splicing.

When the 130 nucleotides long U6 RNA was incubated in the extract, a shorter RNA of approximately the same length as native U6 snRNA (~113 nt) appeared. The analysis of the Uss1p-associated snRNPs revealed that they contained only the shortened U6 RNA. The 130 nt-U6 RNA might have been processed in the yeast extract or it could have been protected from further degradation by the protein(s) interacting with the U6 RNA in the Uss1-containing snRNPs. The length of U6 RNA in other snRNP particles and in the spliceosome has not been investigated; however since the presence of 17 extra nucleotides at the 3' end of U6 RNA did not affect its functionality, it would be highly probable that the shortened U6 RNA (~113 nt) was also contained in other snRNP particles and in the spliceosome. Also, it would have been interesting to investigate the length of the U6 RNA containing six extra nucleotides at its 5' end in snRNP particles. Interestingly, in man, mouse, frog, soya bean and fruitfly cells (Terns *et al.*, 1992; Lund *et al.*, 1992) the U6 snRNA undergoes 3' end posttranscriptional modifications. The U6 snRNA is actually heterogeneous, and diverse RNA species have been distinguished: some U6 snRNAs possessed at their 3' end five oligouridylate residues and a 2'-3'

cyclic phosphate while other U6 snRNAs have a variable length of oligouridylate stretch. In Hela cells, the La and the hnRNP C proteins have been shown to interact only with the stretch of uridylates of the U6 snRNA. The 2'-3' cyclic phosphate end as been shown to be generated within the spliceosome as a consequence of pre-mRNA splicing. These data suggest that the U6 snRNA 3' end modification process is linked with the functionality of the U6 snRNA and/or the formation of snRNP particles (Terns *et al.*, 1992; Tazi *et al.*, 1993; Forné *et al.*, 1995). In yeast cells no 3' end posttranscriptional modification of the U6 snRNA has been reported to date. The analysis on a native gel of the endogenous U6 snRNA in a yeast extract revealed that the U6 snRNA is heterogenous, three U6 snRNA species are sometimes detected in the U6 snRNP particles (M. Cooper, personnal communication). Could the shortening of the 130 nt-U6 RNA to a functional 113 nt-U6 RNA reflect a posttranscriptional U6 snRNA modification? Does the shorter U6 RNA result from a 5' or 3' exonuclease activity or from an endonuclease cleavage? However, this shortening of the 130 nt-U6 RNA may just be an artefact of the *in vitro* system used since the extra 17 nucleotides were added by the T7 RNA polymerase and until now there is no evidence for the transcription *in vivo* of a longer precursor U6 snRNA.

# CHAPTER FIVE

## Detection and Characterisation of Proteins Interacting with the U6 SnRNA

# INTRODUCTION

The mechanistic similarities between the splicing reactions catalyzed by group II introns and eukaryotic spliceosomes favour a possible common evolutionary origin of group II introns and spliceosomes (for review see Cech, 1985; Sharp, 1994). While autocatalytic group II RNAs are able to fold and splice themselves via an extensive structure encoded in the substrate RNA, the eukaryotic pre-mRNAs require for their maturation trans-acting elements (snRNAs and proteins) which interact with one another in a highly ordered fashion in a large complex known as the spliceosome. The spliceosome components as well as transiently associated splicing factors have been suggested to promote the correct folding of the substrate RNA into a splicing-competent configuration and to influence the efficiency and fidelity of the splicing reaction. The study of protein-RNA interactions is therefore very important in order to understand the mechanism of spliceosome assembly and the splicing reaction.

The U6 snRNA undergoes several conformational changes during snRNP interactions and spliceosome assembly. It can be found as a U6 snRNP particle or can be partially base-paired with the U4 snRNA in the U4/U6 snRNP. Once in the spliceosome, U6 snRNA interacts with the pre-mRNA and the U2 snRNA; it is believed to be part of the catalytic centre. At each step of spliceosome assembly numerous well characterised RNA-RNA interactions occur; in contrast much less is known about RNA-protein contacts. The aim of this project is to detect and analyse any U6 snRNA-protein interactions.

In order to study U6 snRNA-protein interactions we have employed a commonly used UV-crosslinking technique. The chemistry of the UV crosslinking is partly understood. Irradiation with UV-light forms covalent linkage between nucleic acid bases and amino acid side-chains. Pyrimidine bases are more photoreactive than purines, and only eleven amino acids appear to be involved in the crosslinking reaction. The most reactive amino acids are cysteine, tyrosine, phenylalanine, histidine, arginine and lysine. The chemical reaction induced by short-wave UV-irradiation is specific: no RNA-RNA crosslinks

occur, nor do proteins become crosslinked to one another (Greenberg, 1979). Proteins and RNA molecules which are close enough one to the other (within one hydrogen bond length) could become crosslinked and can therefore be assumed to have been in direct contact.

The procedure developed to study U6 snRNA-protein interactions and the results obtained are presented in this chapter.

## 5.1 Proteins UV-Crosslinked to U6 SnRNA: The Technique

In a yeast splicing extract U6 snRNPs were reconstituted with a U6 RNA transcribed *in vitro* and radiolabelled at a high specific activity (sections 2.4.4.1 and 2.6.2.3). Typically a 50 µl-100 µl reconstitution reaction was set-up and $^{32}$P- U6 RNA was added at a final concentration of 6 nM. U6 snRNPs were then reconstituted, and spliceosome formation was activated by adding in the reaction unlabelled substrate RNA (1 µM final concentration) transcribed *in vitro* (section 2.4.4.1); water was added instead of the precursor-mRNA when protein-RNA interaction studies were not focused in the spliceosome.

After reconstitution and splicing (section 2.4.4.1), samples were quenched on ice then UV-irradiated (section 2.7) with the presence of 1 µl of nonspecific competitor *E.coli* tRNA solution (10 mg/ml).

After UV irradiation, samples were transferred to a fresh Eppendorf tube and cooled on ice for 5 min. Polyclonal antibodies (Table 10) were used for immunoprecipitation (section 2.8) either under nondenaturing conditions to isolate the entire snRNP particle(s) or the spliceosome, or proteins were denatured prior to immunoprecipitation to isolate only the protein of interest (section 2.8). For the denaturation of proteins prior to the immunoprecipitation, SDS, Triton X-100 and DTT were added at a final concentration of 2% (w/v), 1% (v/v) and 87 mM respectively and the tubes were boiled

for 2 min. An equal volume of 2X IQ buffer was added to the denatured sample and the volume increased to 10 fold with 1X IQ buffer (section 2.8). This mixture was added to the antibodies bound to the PAS beads and the immunoprecipitation was carried out for 2-3 h at 4°C.

The protein complexes attached to the PAS beads were washed 5 times with NTN buffer (section 2.8), then the beads were resuspended in 30 µl of RNase buffer [200 mM NaCl, 20 mM Tris-Cl pH 7.5, 10 mM $MgCl_2$]. Samples were boiled for 1 min in order to destabilize any strong protein-RNA interactions, then ribonucleases T1 and A were added to a final concentration of 7 units/µl and 330 mg/µl respectively. RNase A cleaves RNA molecules adjacent to pyrimidines while the RNase T1 cleaves 3' to G residues only. When a U6 RNA containing at a specific site a photosensitive radiolabelled 4-thioU was used, occasionally the sequence upstream from the labelled site introduced an RNase A or T1 cleavage point which would cleave the radiolabelled phosphate away from the 4-thiouridine. In this case only the RNase which did not cleave this site was used. RNase digestion was carried out for 40 min at 37°C. Then 15 µl of 3X SDS-loading dye [ 94 mM Tris-Cl pH 6.8, 38% (v/v) glycerol, 7.5% (w/v) SDS, 300 mM DTT, 10 mg/ml BPB] were added per sample followed by 10 min boiling. The supernate was split in two and each half was loaded on an SDS-PAGE gel of a different protein resolution ability. Finally, the gel was dried and the proteins UV-crosslinked to the radiolabelled U6 RNA fragments were detected by autoradiography (section 2.2.2).

Typically 3 and 10 day exposures were required to detect proteins UV-crosslinked to uniformly radiolabelled U6 RNA and separated on a 15% and 6% SDS-PAGE gel respectively. When proteins were UV-crosslinked to a 4-thiouridine-containing U6 RNA, an overnight exposure was sufficient to detect proteins.

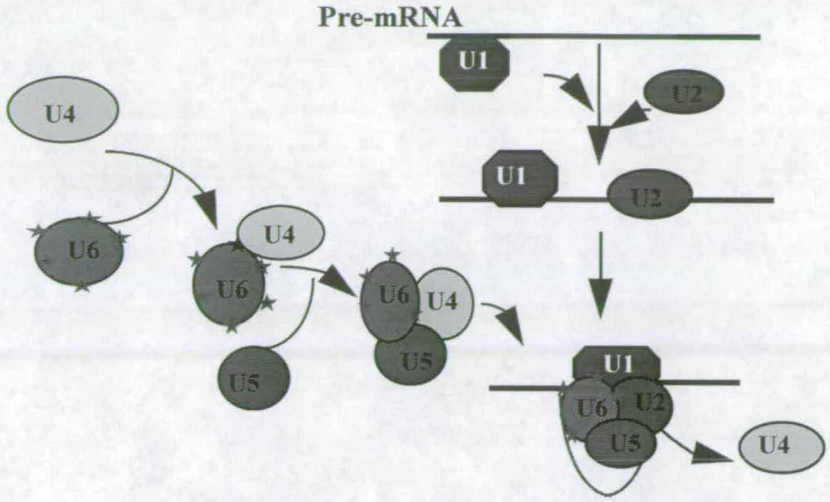
In order to monitor the success of the U6 snRNP reconstitution, splicing activity was assayed. After addition of the synthetic $^{32}$P-U6 RNA, a 9 µl reaction aliquot was transferred to a fresh Eppendorf tube and 1 µl of radiolabelled pre-mRNA (20 000 cpm)

was added. Following incubation at 25°C the RNAs species were recovered (section 2.6.2.2), resolved on a polyacrylamide/urea/TBE gel, and detected by autoradiography (section 2.2.2). Figure 21 summarises the method developed to detect U6 snRNA-protein interactions.

## 5.2 Studies of Prp24p, Uss1p and Prp8p Contacts with the U6 SnRNA

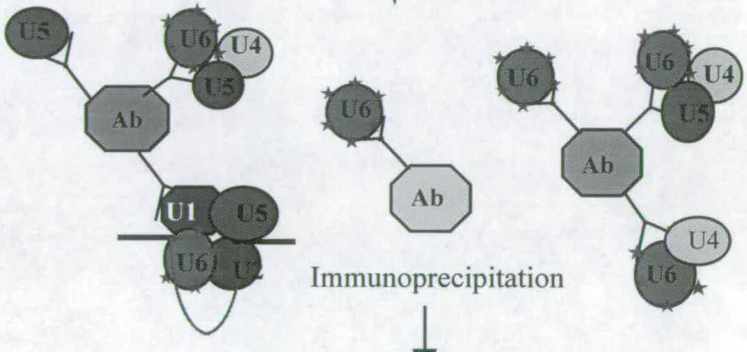
In order to detect protein-U6 RNA interactions the experimental protocol described above was followed (section 5.1) and anti-Prp24, anti-Uss1 and anti-Prp8 antisera (Table 10) were used to immunoprecipitate snRNP particles.

Uss1p (21 kD) is associated with the U6 snRNA (Cooper *et al.*, 1995). Crude anti-Uss1 antibodies co-immunoprecipitate U6 snRNP, U4/U6 snRNP and to a lesser extent U4/U6.U5 tri-snRNP (Cooper *et al.*, 1995). Prp24p (45-50 kD) is a U6 snRNP associated protein which contacts the U6 snRNA directly (Ghetti *et al.*, 1995; Jandrositz and Guthrie, 1995). The two batches (1015 and 1016) of anti-Prp24 antisera prepared in this work (Chapter 3) were tested for their ability to precipitate Prp24p UV-crosslinked to the U6 RNA. Polyclonal anti-Prp24 antisera co-immunoprecipitate the U6 snRNP (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995; section 3.3.3). Prp8p is a 280 kD U5 snRNP protein which also associates with the tri-snRNP (Lossky *et al.*, 1987). Prp8p is required for the stability of the tri-snRNP and its association into the spliceosome (Brown *et al.*, 1992). Prp8p remains in the spliceosome and interacts at least with the U5 snRNA and the pre-mRNA (Teigelkamp *et al.*, 1995). The anti-Prp8 antibodies (Table 10) used in these experiments are co-immunoprecipitating the U5 snRNP, the tri-snRNP and the entire spliceosome.




*In vitro* U6 snRNP Reconstitution with $^{32}$P-U6 RNA and Splicing Reaction

UV-Irradiation

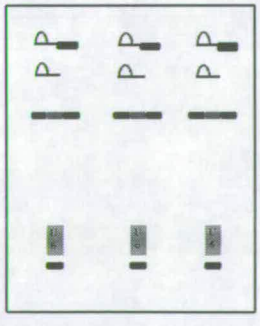



Aliquot of the U6 snRNP Reconstitution Mixture

$^{32}$P pre-mRNA

Splicing reaction

RNA Recovery and Resolution on a polyacrylamide/urea Gel

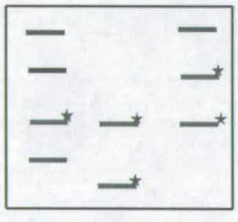


Detection of the RNA by Autoradiography

Immunoprecipitation

RNase T1 and/or A Treatment

Protein Separation on an SDS-Polyacrylamide Gel



Proteins-UVcrosslinked to $^{32}$P-U6 RNA Detected by Autoradiography

### 5.2.1 Preliminary characterisation of the different tools required to detect U6 RNA-protein interactions

#### 5.2.1.1 Analysis of the two batches of crude anti-Prp24 antisera produced from rabbits

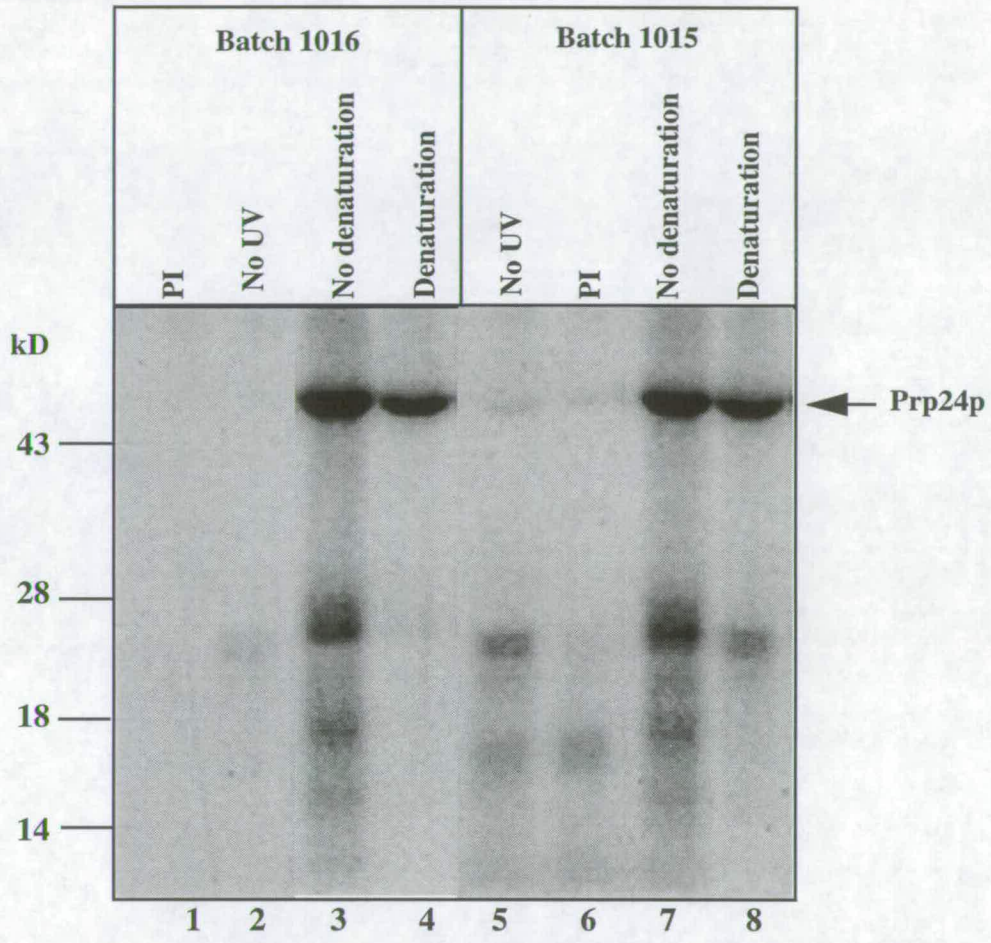
Both batches (1015 and 1016) of anti-Prp24 antisera were tested for their ability to immunoprecipitate Prp24p UV-crosslinked to the radiolabelled U6 RNA ($^{32}$P-αATP) (section 5.2) (Figure 22). In the absence of UV-irradiation, two rather diffuse signals (~15 kD and 20 kD) were detected with antiserum batch 1015 (lane 5); only the 20 kD band was observed with batch 1016 (lane 2). A weak signal at 15 kD was noticed with the preimmune serum of batch 1015 (lane 6) but no signal was detected with that of batch 1016 (lane 1). The 15 kD and the 20 kD signals were neither UV-crosslinking nor antibody specific thus they were not of interest. Following UV irradiation a protein of ~45 kD contacted the U6 RNA and was immunoprecipitated with both antisera under nondenaturing conditions (lanes 3 and 7) or after protein denaturation (lanes 4 and 8), indicating that Prp24p was UV-crosslinked to the U6 RNA. When the crude antiserum batch 1015 was used to immunoprecipitate the denatured protein, in addition to Prp24p another nonspecific signal (~20 kD) was also detected. However only the UV-crosslinked Prp24p was precipitated with antiserum batch 1016. Two other proteins (~16 kD and 25 kD) were also bound to the U6 RNA and co-immunoprecipitated with both batches of antisera under nondenaturing conditions (lanes 3 and 7). Under these particular experimental conditions (i.e. U6 RNA was radiolabelled with $^{32}$P-αATP), no proteins of a high molecular weight were UV-crosslinked to the U6 RNA, immunoprecipitated with anti-Prp24 antibodies and detected on a 6% SDS-polyacrylamide gel (data not shown).

The analysis of both anti-Prp24 antisera demonstrated that less background signal was detected with batch 1016 which was therefore more suitable to investigate U6 RNA-protein contacts.

**Figure 22. Proteins UV-crosslinked to the U6 RNA and immunoprecipitated with crude anti-Prp24 antisera.**

U6 snRNPs were reconstituted (50 µl reaction) with a U6 RNA transcribed *in vitro* using pUC12U6 linearized with *Dra*I as template and $^{32}$P-αATP as source of radioactivity (section 2.4.4.1). Proteins were UV-crosslinked to the nucleic acids and then incubated with preimmune serum (lanes 1 and 6) or with the crude anti-Prp24 antiserum under nondenaturing conditions (lane 3 and 7) or after protein denaturation (lanes 4 and 8). UV-irradiation was omitted in lanes 2 and 5. Immunoprecipitation and washes were carried out with 150 mM NaCl. Proteins were resolved on a 15% SDS-polyacrylamide gel and detected by autoradiography.

Anti-Prp24 Antisera




15% SDS-PAGE

5.2.1.2 Characterisation of affinity purified anti-Prp24 and anti-Uss1 antibodies

The analysis by Western blotting of proteins detected from a total yeast extract with the crude anti-Prp24 or anti-Uss1 antisera indicated that both lots of antibodies were reacting with Prp24 or Uss1 respectively and also with numerous additional proteins, see Figure 23 (Panels A and B, lane 1).

Nonspecific UV-crosslinking signals were detected when both batches of crude anti-Prp24 antisera were used to immunoprecipitate the protein complexes, with batch 1016 being better (Figure 22). Also when denatured proteins UV-crosslinked to U6 RNA were immunoprecipitated with the crude anti-Uss1 antiserum, Uss1p and two additional polypeptides (~45 kD and ~190 kD) were detected (data not shown). To improve the specificity of detection of the proteins UV-crosslinked to the U6 RNA, anti-Prp24 and anti-Uss1 antibodies were affinity purified from their respective crude serum (section 2.9.5). Also the use of affinity purified anti-Uss1 antibodies would indicate whether the additional polypeptides (~45 kD and 190 kD) UV-crosslinked to the U6 RNA and immunoprecipitated after proteins denaturation were contaminating proteins or if they were sharing homology with Uss1p. No background signal was detected when the crude anti-Prp8 antiserum was used in the immunoprecipitation step; therefore no affinity purification of these antibodies was required.

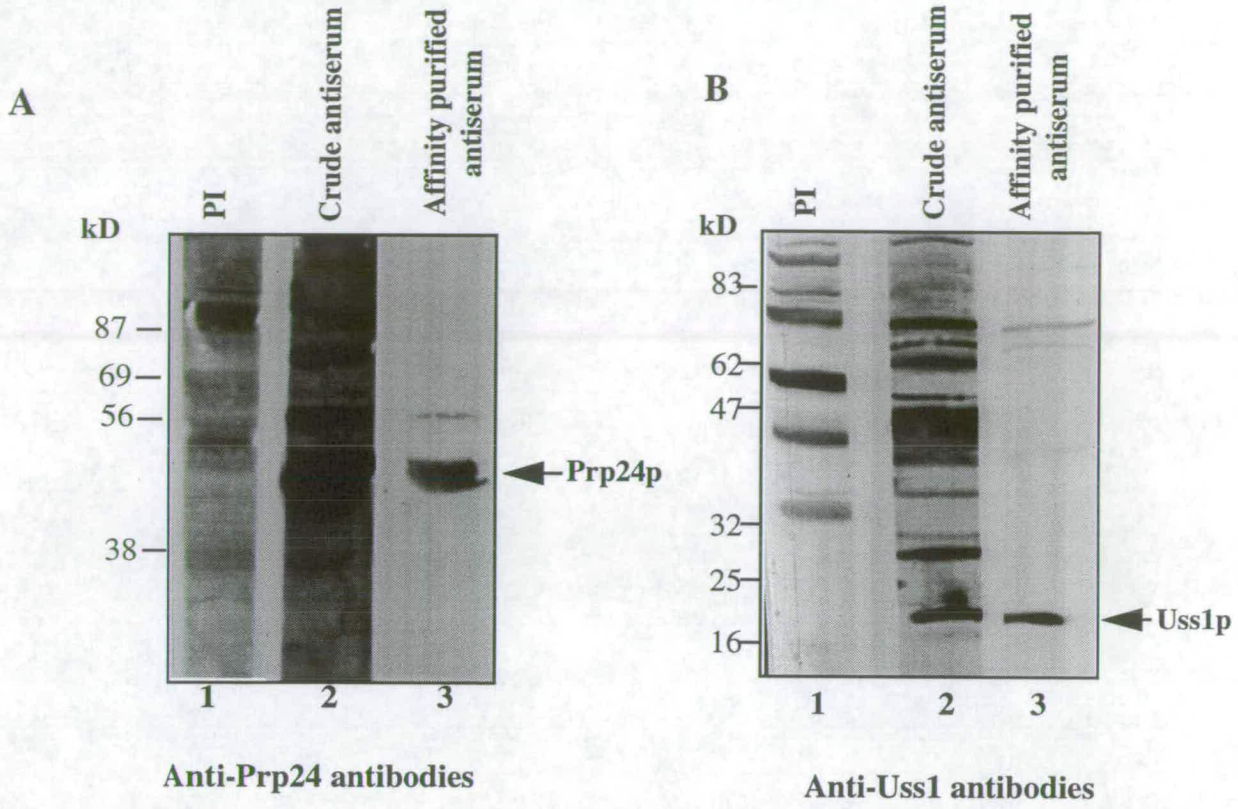
In order to limit the cross reactivity with other unrelated polypeptides, anti-Prp24 antibodies were affinity purified (section 2.9.5 ) against a fusion protein (6X His-Prp24) different from the one (β-galactosidase-Prp24p) used to raise polyclonal antibodies from rabbits (Table 10). Likewise anti-Uss1 antibodies were affinity purified against a (TrpE-Uss1) fusion protein while they were raised against Uss1p fused to the β-galactosidase (Table 10). The proteins detected by Western blotting with the crude or the purified anti-Prp24 and anti-Uss1 antibodies are shown in Figure 23. The affinity purification of both antisera greatly decreased the number of cross-reacting proteins (Panels A and B, compare lanes 2 and 3). In addition to Prp24p, one weaker signal (~60 kD) was detected from a yeast total extract with the affinity purified anti-Prp24

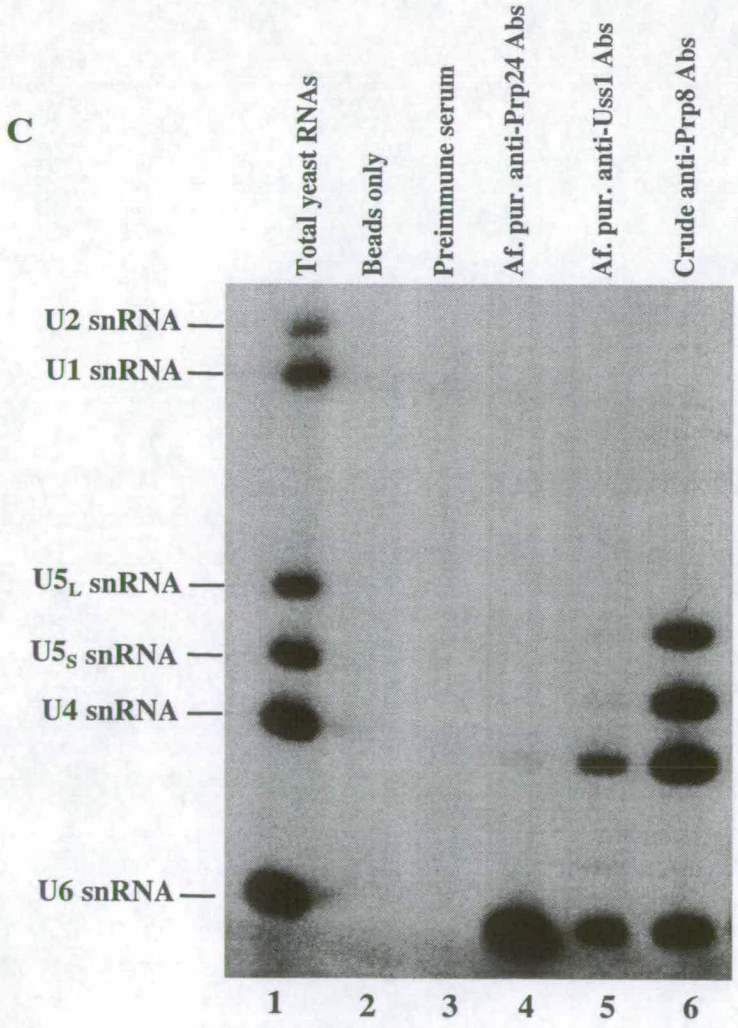
**Figure 23. Characterisation of the affinity purified anti-Prp24 and anti-Uss1 antibodies.**

**A:** Detection by Western blotting of Prp24p and Uss1p. Yeast S150-2B (pBM24:*CEN4*, *URA3*, *GAL1-PRP24*) cells (Tables 7 and 9) were grown in YMGlyLac medium (section 2.1.2.3) supplemented with tryptophan (Table 5) and galactose was added (20% (v/v) final) when cells reached the exponential phase. After 5h galactose-induction of *PRP24* expression, cells were harvested and proteins extracted (section 2.5.2). The protein content was estimated by Bradford assay (section 2.5.1) and proteins were loaded on a 10% SDS-polyacrylamide gel. After resolution, proteins were transferred to a PVDF membrane by Western blotting (section 2.5.6), detected by Ponceau staining and the membrane was cut into several bands each containing an equal amount of proteins (100 µg). A band was incubated with the preimmune serum 1/1000 fold diluted (lane 1) or with the crude anti-Prp24 antiserum batch 1016 (1/1000 fold dilution) (lane 2) or with the affinity purified anti-PRP24 antibodies (1/1000 fold dilution) (lane 3).

**B:** Yeast MCY4 (*LEU2-GAL1-USS1*, pYEP24-USS1) cells (Tables 7 and 9) were grown in YPDA (section 2.1.2.3), cells were harvested at the exponential phase and proteins extracted (section 2.5.2). Proteins were resolved on a 15% SDS-polyacrylamide gel, transferred on a PVDF membrane which was cut into several bands (each containing 100 µg of proteins). One band was incubated with the preimmune serum (1/1000 fold dilution) (lane 1), or with the crude anti-Uss1 antiserum (1/1000 fold dilution) (lane 2) or with the affinity purified anti-Uss1 antibodies (lane 3).

**C:** Sn-RNAs associated with the affinity purified anti-Prp24 and anti-Uss1 antibodies and with the crude anti-Prp8 antiserum. A yeast splicing extract (40 µl) prepared from SC261 cells (Table 7) was added in a splicing reaction (100 µl) and snRNAs were co-immunoprecipitated with the affinity purified anti-Prp24 (lane 4) or anti-Uss1 (lane 5) antibodies or with the crude anti-Prp8 antiserum (lane 6). As controls, the mixture was incubated with PAS beads only (lane 2) or with the preimmune serum (lane 3). Immunoprecipitation and washes (section 2.8) were carried out under 150 mM NaCl. Co-immunoprecipitated snRNAs were recovered, resolved on a 6% urea/polyacrylamide gel and tranferred on a nylon membrane (section 2.4.4.3). The presence of U1, U2, U4, U5 and U6 snRNAs was detected by hybridization with end-radiolabelled (section 2.4.3.10) oligodeoxynucleotides 486A, G1803, 483A, 485A and Taq6A respectively (Table 8). Total yeast RNA was extracted from 5 µl of splicing extract (section 2.8) (lane 1). SnRNAs were detected by autoradiography.


A
PI
Crude antiserum
Affinity purified antiserum
kD
87
69
56
38
Prp24p
1
2
3
Anti-Prp24 antibodies
B
PI
Crude antiserum
Affinity purified antiserum
kD
83
62
47
32
25
16
Uss1p
1
2
3
Anti-Uss1 antibodies


C
Total yeast RNAs
Beads only
Preimmune serum
Af. pur. anti-Prp24 Abs
Af. pur. anti-Uss1 Abs
Crude anti-Prp8 Abs
U2 snRNA
U1 snRNA
U5L snRNA
U5S snRNA
U4 snRNA
U6 snRNA
1
2
3
4
5
6

antibodies. Also, in addition to Uss1p, two other weaker signals (a doublet at ~70-80 kD) were detected with the affinity purified anti-Uss1 antibodies.

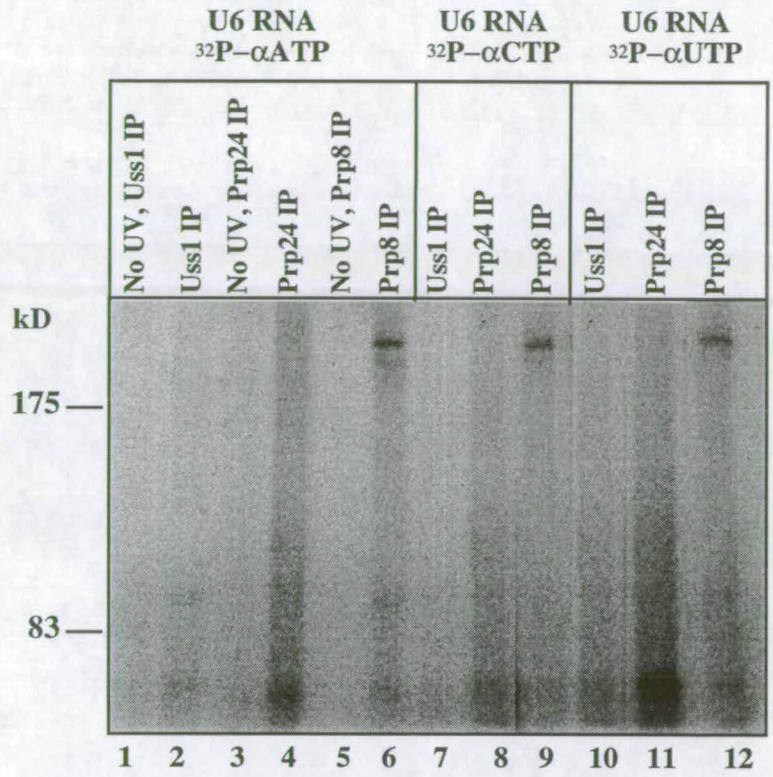
The ability of the affinity purified anti-Uss1 and anti-Prp24 antisera to co-immunoprecipitate snRNAs was tested (section 2.8) and compared with one another and with the snRNAs co-immunoprecipitated with anti-Prp8 antibodies (Figure 23; Panel C). Like the crude anti-Prp24 antiserum (see Figure 17), the affinity purified antibodies co-immunoprecipitated the U6 snRNA and very little of U4 snRNA (lane 4). The affinity purified anti-Uss1 antibodies co-immunoprecipitated the U4 snRNA and the U6 snRNA, no U5 snRNA was detected (lane 5), whereas a small fraction of U5 snRNAs was co-immunoprecipitated with the crude antiserum (Cooper *et al.*, 1995, and see Figure 35). The U4 snRNA, the U6 snRNA and the two species of U5 snRNAs were co-immunoprecipitated with anti-Prp8 antibodies (lane 6). No U1 or U2 snRNAs were detected since no precursor mRNA was added in the reaction mix, therefore very little, if any, spliceosomes formed and were isolated with anti-Prp8 antibodies.

#### 5.2.1.3 Influence of the radiolabelled U6 RNA on the U6 RNA-protein contacts detected

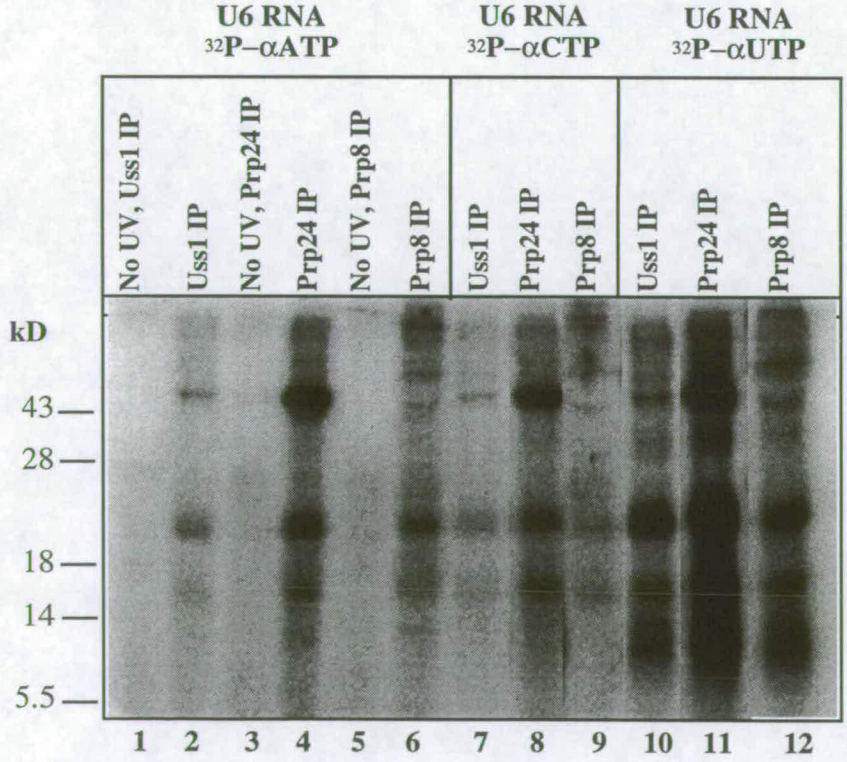
The DNA template (PCU6) containing 6 extra nucleotides at its 5' end (see section 4.4) was used to transcribe *in vitro* a U6 RNA radiolabelled with either $^{32}$P-αATP, $^{32}$P-αCTP or $^{32}$P-αUTP. U6 snRNPs were reconstituted with each of these U6 RNAs containing the same specific activity. Subsequently, RNA-protein contacts were investigated by UV-crosslinking and protein complexes were immunoprecipitated with anti-Prp24, anti-Uss1 and anti-Prp8 antibodies. The signals detected with U6 RNA labelled with $^{32}$P-αATP are UV-irradiation dependent (Figure 24, lanes 1, 3, 5). The digestion with RNase A and T1 of U6 RNA produces fragments which contain more A than T or C residues, since neither of these enzymes cleaves after an A. Therefore, even if the UV-crosslinking dependency of the signals has not been investigated with U6 RNA labelled with $^{32}$P-αCTP or with $^{32}$P-αUTP, it is more likely that no signal would

**Figure 24. Proteins UV-crosslinked to U6 RNA randomly radiolabelled with different nucleotides.**

U6 RNA was transcribed from a DNA template amplified by PCR (PCU6) with oligodeoxynucleotides Lory1 and S5761 (Table 8) used as primers. This U6 RNA was radiolabelled with either $^{32}$P-αATP, $^{32}$P-αCTP or $^{32}$P-αUTP with the same specific activity (section 2.4.4.1). U6 snRNPs were reconstituted (100 µl) with 5 nM of each of these U6 RNAs. Proteins were UV-crosslinked to the nucleic acids and co-immunoprecipitated with the affinity purified anti-Prp24 (lanes 4, 8, 11) or anti-Uss1 (lanes 2, 7, 10) antibodies or with the crude anti-Prp8 antiserum (lanes 6, 9, 12). UV-crosslinking was omitted in samples 1, 3 and 5. Protein complexes were washed with NTN buffer (section 2.8) containing 150 mM NaCl, half of the proteins were loaded on a 6% SDS-polyacrylamide gel and the other half were loaded on a 15% SDS-polyacrylamide gel.




6% SDS-PAGE




15% SDS-PAGE

have been observed in the absence of UV-irradiation. After U6 snRNP reconstitution and UV-crosslinking, the intensity of the bands varies according to the radioactive nucleotide used to label the U6 RNA. For each radioactive U6 RNA tested, proteins were most efficiently co-immunoprecipitated with anti-Prp24 antibodies. Weaker signals were obtained with the U6 RNA labelled with $^{32}$P-αATP (lanes 2, 4, 6) or with $^{32}$P-αCTP (lanes 7-9) compared with $^{32}$P-αUTP, which gave by far the strongest UV-crosslinking signals (lanes 10-12). Also with $^{32}$P-αUTP, an additional 10 kD protein was UV-crosslinked and immunoprecipitated with all the antisera used.

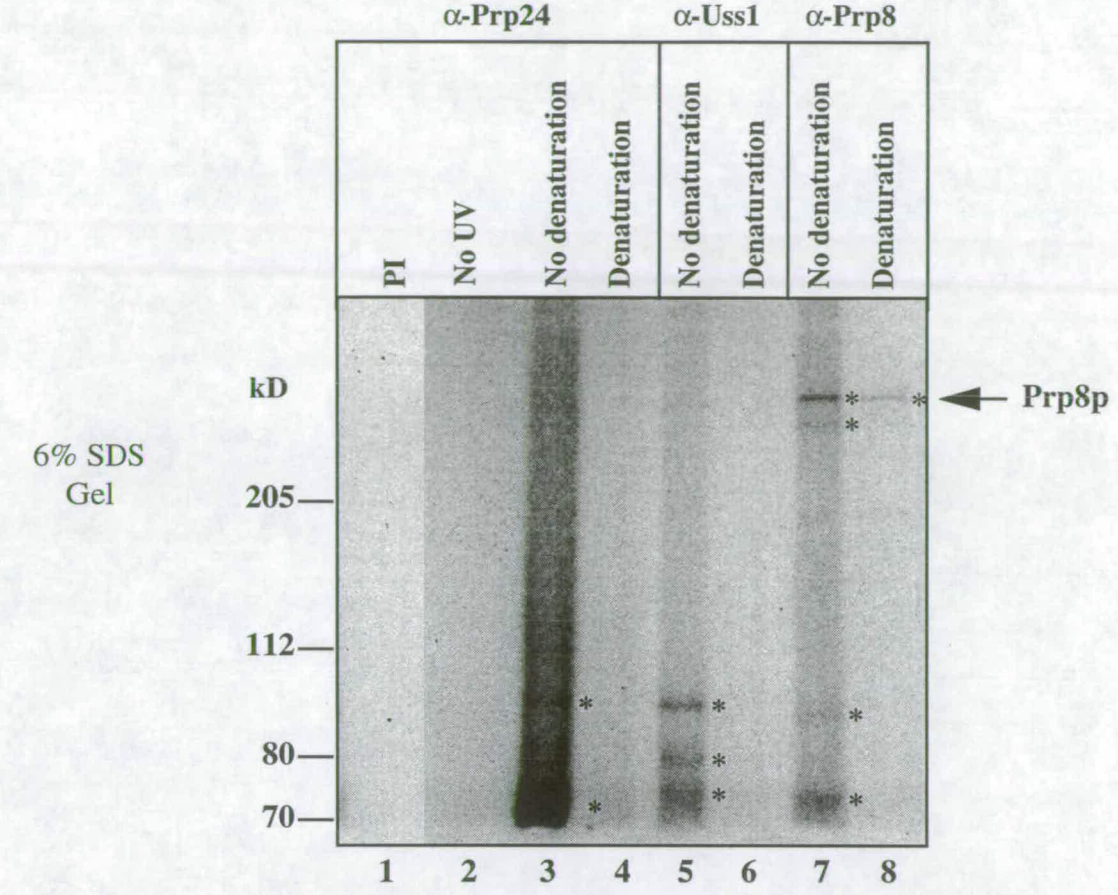
The use of $^{32}$P-αUTP to radiolabel the U6 RNA transcribed *in vitro* from PCU6 template, gave the best detection of proteins UV-crosslinked to the RNA. Therefore U6 RNA was radiolabelled only with $^{32}$P-αUTP while the U6 RNA transcribed from pUC12U6 linearized with *Dra*I was radiolabelled with $^{32}$P-αATP. In order to distinguish between these two RNAs, we conventionally named the former U6 RNA ($^{32}$P-αUTP) and the latter U6 RNA ($^{32}$P-αATP).

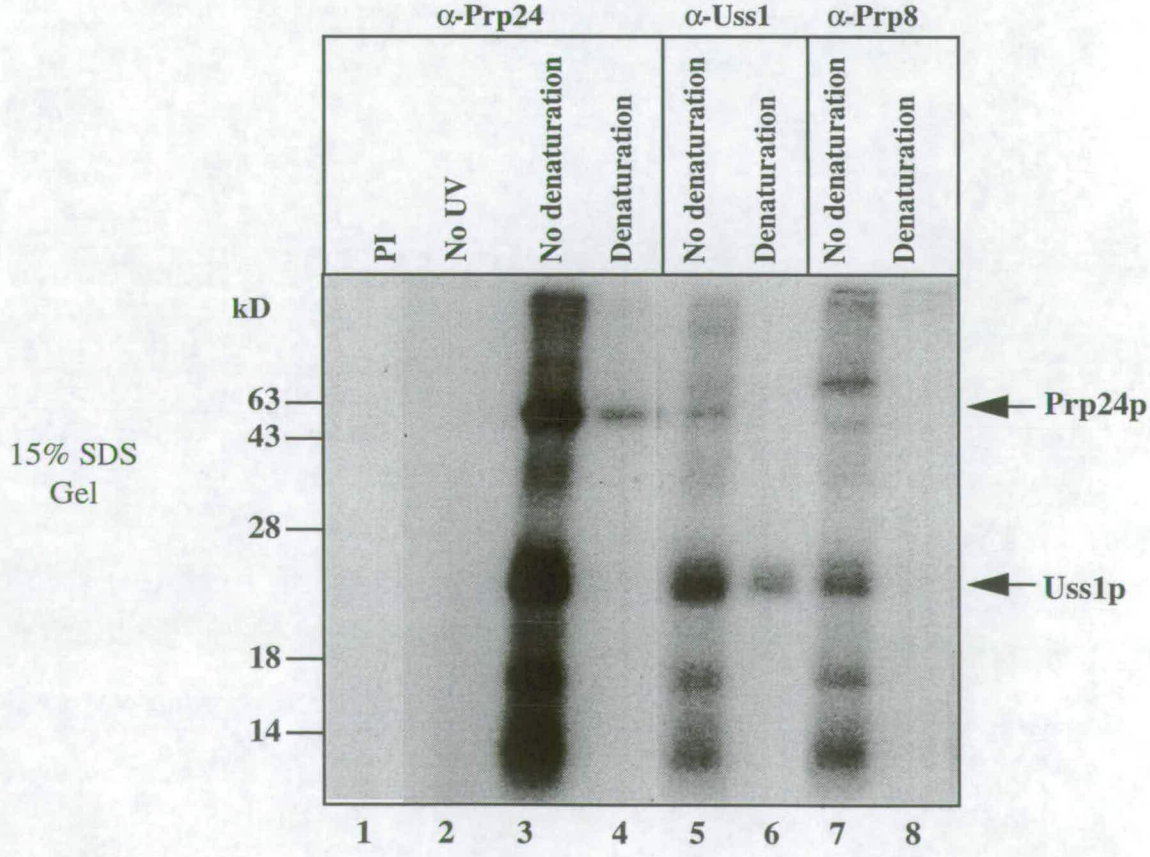
### 5.2.2 Proteins contacting the U6 RNA and isolated with anti-Prp24, anti-Uss1 and anti-Prp8 antibodies

Immunoprecipitated proteins detected after U6 snRNP reconstitution with the U6 RNA ($^{32}$P-αUTP) and UV-crosslinking are shown in Figure 25. Prp24p contacted the U6 RNA and was immunoprecipitated with the affinity purified anti-Prp24 antibodies under nondenaturing conditions or after protein denaturation (15% gel; lanes 3, 4). This binding was antiserum and UV-irradiation specific (lanes 1 and 2). A protein of ~25 kD contacted the U6 RNA and was immunoprecipitated with the affinity purified anti-Uss1 antibodies under nondenaturing conditions or after protein denaturation (15% gel; lanes 5 and 6), thus Uss1p contacted the U6 RNA. Likewise a protein of > 210 kD was detected after UV-crosslinking and immunoprecipitation with anti-Prp8 antibodies under nondenaturing conditions or after protein denaturation (6% gel; lanes 7 and 8) indicating that Prp8p was also in close proximity to the U6 RNA.

**Figure 25. Proteins UV-crosslinked to the randomly radiolabelled U6 RNA ($^{32}$P-αUTP).**

A U6 snRNP reconstitution reaction (100 µl) was prepared, proteins were immunoprecipitated with the affinity purified anti-Prp24 antibodies under nondenaturing conditions (lane 3) or after proteins denaturation (lane 4), or anti-Uss1 antibodies under nondenaturing conditions (lane 5) or after protein denaturation (lane 6). UV-crosslinked proteins were also co-immunoprecipitated with the crude anti-Prp8 antiserum under nondenaturing conditions (lane 7) or after protein denaturation (lane 8). Samples were incubated with the preimmune serum in lane 1 or were not UV-irradiated in lane 2. Precipitations and washes were carried out under 150 mM NaCl. Half of the protein complexes were resolved on a 6% SDS-polyacrylamide gel and the other half on a 15% SDS-polyacrylamide gel. The star symbol on the 6% SDS-PAGE highlight the protein bands.


α-Prp24
α-Uss1
α-Prp8
PI
No UV
No denaturation
Denaturation
No denaturation
Denaturation
No denaturation
Denaturation
kD
6% SDS Gel
205
112
80
70
Prp8p
1 2 3 4 5 6 7 8


α-Prp24
α-Uss1
α-Prp8
PI
No UV
No denaturation
Denaturation
No denaturation
Denaturation
No denaturation
Denaturation
kD
15% SDS Gel
63
43
28
18
14
Prp24p
Uss1p
1 2 3 4 5 6 7 8

On a 6% and 15% SDS-PAGE gels (Figure 25), five additional proteins (10 kD, 16 kD, 25 kD, 70-75 kD and 90 kD; lane 3) were co-immunoprecipitated under nondenaturing conditions with anti-Prp24 antibodies. Six other polypeptides (10 kD, 16 kD, 45 kD, 70-75 kD, 80 kD and 90 kD; lane 5) were UV-crosslinked to U6 RNA and co-immunoprecipitated with anti-Uss1 antibodies. At least eight other polypeptides (10 kD, 16 kD, 25 kD, 43 kD, 65-70 kD, 70-75 kD and ~85 kD; one at >210 kD; lane 7) were detected with anti-Prp8 antiserum under nondenaturing conditions.

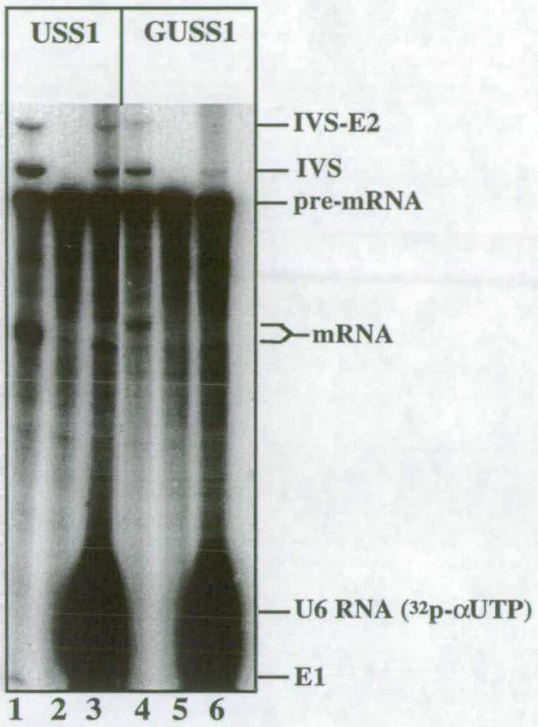
Proteins of similar sizes were UV-crosslinked to the U6 RNA and co-immunoprecipitated with the three different antisera tested. In particular, the 25 kD protein detected with anti-Prp24 and anti-Prp8 antibodies migrated to the same position as Uss1p. In order to identify this polypeptide we followed Uss1p in snRNP particles immunoprecipitated with the three lots of antibodies previously used. For this purpose, splicing extract was prepared from MCY4 (*LEU2-GAL1-USS1*, pAEM13; Tables 7 and 9) cells grown on glucose-containing medium. In this extract cells overproduce a Gal4-Uss1 fusion protein (Guss1p) as only the source of Uss1p. This fusion protein possesses a theoretical molecular weight of 37 kD and is functional since it complements the Uss1p defect *in vivo* (A. Mayes, this lab; personal communication). The splicing activity and the U6 snRNP reconstitution efficiency of an extract containing only Guss1p were tested (Figure 26; Panel A). The splicing activity of the extract prepared from MCY4 (pAEM13) was weaker than that detected in an extract prepared from wild-type SC261 cells (Table 7) (compare lane 1 versus 4). The spliced RNA and the lariat-intermediate detected in the former extract were longer that the one observed in the latter. Maybe in MCY4 (pAEM13) cells the pre-mRNAs were polyadenylated. No splicing activity was detected when the endogenous U6 snRNA was degraded (lanes 2 and 5). Splicing activity was recovered in both extracts by addition of U6 RNA ($^{32}$P - αUTP) (lanes 3 and 6); about 50 % of the initial activity (in respect to excised lariat-intron, and lariat-intron-exon2 signals) was restored after U6 snRNP reconstitution. This demonstrated that the Guss1 fusion protein was also functional *in vitro* and had no detrimental effect on the reconstitution of the U6 snRNP particles and the splicing activity of the extract.

**Figure 26. Uss1p contacts the U6 RNA in the U6 snRNP, the U4/U6 snRNP and in the tri-snRNP.**

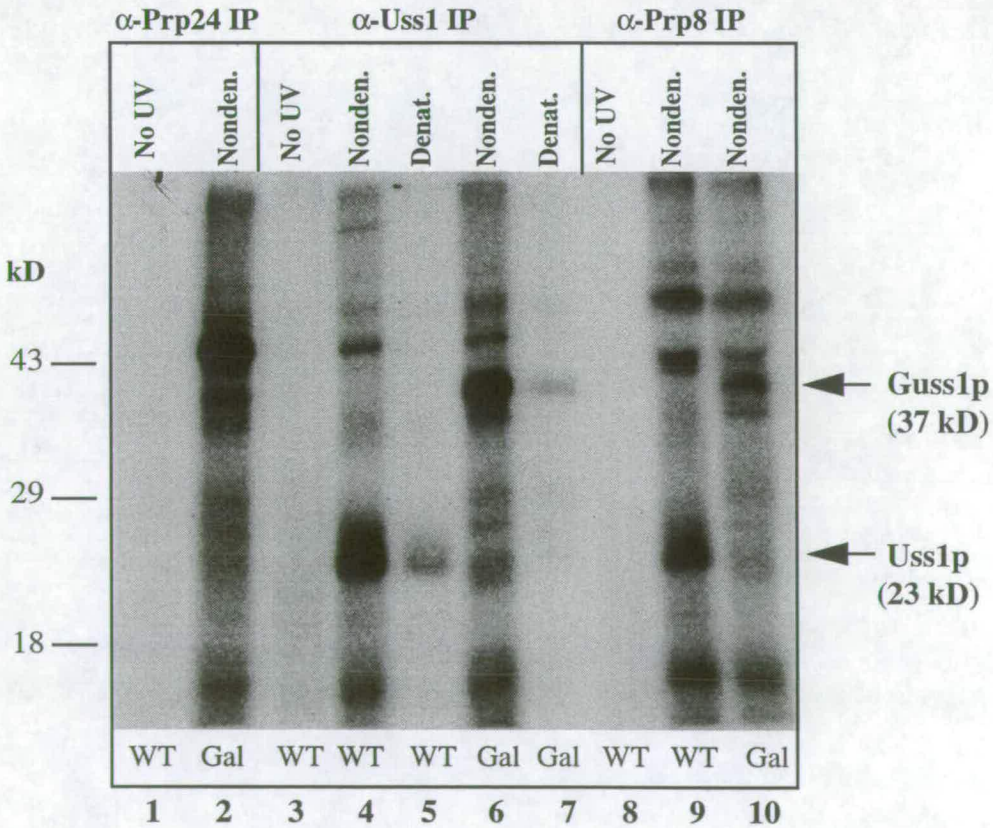
**A:** U6 snRNP reconstitution in extracts producing either Uss1p or Guss1p. The initial splicing activity of an extract prepared from SC261 cell (Lane 1) or from MCY4 (pAEM13) cells grown on YPDA (Tables 7 and 9) (lane 4) was tested after preincubation at 30°C for 30 min. The addition of 140 nM of U6 knockout DNA oligo was sufficient to completely deplete the endogenous U6 snRNA and abolish the splicing activity of the extracts (lanes 2 and 5). U6 snRNPs were reconstituted (60 µl reaction) with 5 nM of randomly radiolabelled U6 RNA ($^{32}$P-αUTP). An aliquot (9 µl) was taken off and 1 µl of radiolabelled precursor mRNA (20 000 cpm) was added (lanes 3 and 6). After splicing, RNA species were extracted, loaded on a 7M urea/polyacrylamide gel and detected by autoradiography. IVS: excised lariat-intron; IVS-E2: lariat-intron-exon2; E1: exon1; pre-mRNA: actin substrate RNA (p283, Table 9); mRNA: spliced RNA; U6 RNA: RNA produced by transcription *in vitro* and radiolabelled with $^{32}$P-αUTP.

**B:** Investigation of U6 RNA-protein contacts in U6 snRNP, U4/U6 snRNP and U4/U6.U5 tri-snRNP. U6 snRNPs (60 µl reaction) were reconstituted with 5 nM of U6 RNA ($^{32}$P-αUTP) after UV-crosslinking proteins were immunoprecipitated with affinity purified anti-Prp24 antibodies under nondenaturing conditions (lane 2) or with affinity purified anti-Uss1 antibodies under nondenaturing conditions (lanes 4 and 6) or after protein denaturation (lanes 5 and 7). Proteins complexes were also immunoprecipitated with the crude anti-Prp8 antiserum under nondenaturing conditions (lanes 9 and 10). UV-irradiation was omitted in lanes 1, 3, 5 and 8. Protein complexes were washed with NTN buffer (section 2.8) containing 150 mM NaCl, loaded and resolved on a 15% SDS-polyacrylamide gel. WT: splicing extract prepared from wild-type SC261 cells; Gal: splicing extract prepared from MCY4 (pAEM13) cells.

A




B




15% SDS-PAGE

U6 snRNPs were reconstituted with U6 RNA ($^{32}$P-αUTP) from a wild-type extract or from an extract containing the Guss1 fusion protein. After UV-crosslinking proteins were immunoprecipitated with different antisera and resolved on a 15% SDS-PAGE (Figure 26, Panel B). Little difference was observed in the intensity of the signals between the two extracts tested (compare lanes 4/6, 5/7 and 9/10, and compare Figure 25 lane 3 with Figure 26 lane 2). Therefore the weaker splicing activity of the extract containing Guss1p did not affect the U6 RNA-protein contacts. The reconstitution efficiency of the U6 snRNP particles was probably similar in both extracts. In extract prepared from MCY4 (pAEM13) grown on YPDA (section 2.1.2.3), signals in the 20-30 kD range were much reduced whereas a protein of a size (~40 kD) similar to Guss1p was observed which was not detected in a wild-type extract (compare lanes 4 and 6). This 40 kD polypeptide was also immunoprecipitated with anti-Uss1 antibodies after protein denaturation (lane 7) demonstrating that Guss1p contacted the U6 RNA. The fusion of the Gal4 to Uss1p did not interfere with its interaction with the U6 RNA. Also in the extract containing the fusion protein, two additional proteins (16 kD and 45 kD) were UV-crosslinked to the U6 RNA and immunoprecipitated under nondenaturing conditions with anti-Uss1 antibodies (lane 6). Similar signals were detected when protein extracted from SC261 cells were UV-crosslinked to the U6 RNA and immunoprecipitated with the same antibodies (lane 4) suggesting that, if these bands corresponded to identical polypeptides, their binding to the U6 molecule was extract independent. The 25 kD protein UV-crosslinked to the U6 RNA from a wild-type extract (Figure 25, 15 % gel; lane 3) and co-immunoprecipitated with anti-Prp24 antibodies was not detected anymore in an extract containing the fusion Guss1 protein, instead, a new protein (~40 kD) was observed (Fig 26, lane 2). Therefore the 25 kD protein was indeed Uss1p in close proximity to the U6 RNA in a Prp24p-containing protein complex. Likewise, the 25 kD protein UV-crosslinked to the U6 RNA and co-immunoprecipitated with anti-Prp8 antibodies was shifted up to a band of ~40 kD in an extract containing Guss1p (Figure 26, compare lanes 9 and 10); thus Uss1p also contacted the U6 RNA in Prp8p-containing protein complex.

By using the same band shift principle it was attempted to identify the 45 kD polypeptide co-immunoprecipitated with anti-Uss1 antisera. As the Prp24 protein tagged with the FLAG-epitope was not functional (Chapter 3), YVDP24 (*prp24::HIS3/PRP24*) cells were transformed with the plasmid pGPRP24 (Table 9). This plasmid harbours the gene encoding Prp24p fused at its N-terminus with the Gal4 DNA binding domain (A. Colley, this lab). The corresponding fusion protein (Gprp24p) possessed a theoretical molecular weight of 65-70 kD. Despite several trials, no sporulation of YVDP24 (pGPRP24) was obtained, therefore no functional complementation test was carried out for this fusion protein. Splicing extracts were prepared from YVDP24 (pGPRP24) and proteins were UV-crosslinked to the U6 RNA, denatured and immunoprecipitated with anti-Prp24 antibodies. Two protein signals (45 kD, ~65 kD) were detected (data not shown) with the 65 kD band being much weaker than Prp24p signal. This suggested that compared to the wild-type protein, much less of the fusion Gprp24p was incorporated into snRNP particles and/or UV-crosslinked to the U6 RNA. As the signal for the 45 kD protein co-immunoprecipitated with anti-Uss1 antibodies was already weak (Figure 26, compare the 45 kD band in lanes 2 and 4), the inefficiency of the UV-crosslinking of the fusion Gprp24 protein prevented its use for a band shift assay.

Protein-U6 RNA contacts reported in this section are summarised in Table 12.

| U6 snRNP (affinity purified anti-Prp24 antisera) | | U6 snRNP + U4/U6 snRNP (affinity purified anti-Uss1 antibodies) | | U5 snRNP + U4/U6.U5 tri-snRNP (crude anti-Prp8 antisera) | |
|---|---|---|---|---|---|
| **15% Gel** | **6% Gel** | **15% Gel** | **6% Gel** | **15% Gel** | **6% Gel** |
| | | | | | > 210 kD (Prp8p) |
| | | | | | > 210 kD |
| | ~90 kD | | ~90 kD | | |
| | | | | | ~85 kD |
| | | | 80 kD | | |
| | 70-75 kD | | 70-75 kD | | 70-75 kD |
| | | | | 65-70 kD | |
| 45 kD (Prp24p) | | 45 kD | | | |
| | | | | 43 kD | |
| 25 kD (Uss1p) | | 25 kD (Uss1p) | | 25 kD (Uss1p) | |
| 16 kD | | 16 kD | | 16 kD | |
| 10 kD | | 10 kD | | 10 kD | |

Protein molecular weights indicated in Table 12 were determined according to the molecular markers run alongside the reaction samples on the same protein gel. The size

values reported are not very precise and depend on the accuracy of the supplied protein ladder and the resolution ability of the protein gel used.

## 5.3 Discussion

Western blotting analysis indicated that both affinity purified anti-Prp24 and anti-Uss1 antibodies recognised specifically Prp24p and Uss1p in a total protein extract prepared from yeast cells overproducing Prp24p or Uss1p respectively. One additional polypeptides (65 kD) was detected with anti-Prp24 antisera; also two other proteins (~70-80 kD) were recognised by anti-Uss1 antisera. These proteins were most likely cross-reacting polypeptides. The analysis by Northern blotting of the snRNAs co-immunoprecipitated with the affinity purified antisera indicated that the purification of anti-Prp24 antibodies had not altered their ability to precipitate the U6 snRNA and traces of U4 snRNA. The affinity purified anti-Uss1 antibodies co-immunoprecipitated the U4 snRNA and the U6 snRNA but unlike the crude antiserum, no U5 snRNA was recovered (see Figure 35 and Cooper *et al.*, 1995). Probably the antibodies recognising Uss1p in the tri-snRNP have not been selected during the purification procedure. The affinity purification of both lots of antibodies greatly improved their binding specificity since in the absence of UV-irradiation no proteins were detected after immunoprecipitation. When denatured proteins were immunoprecipitated with the affinity purified anti-Prp24 or anti-Uss1 antibodies only Prp24 and Uss1 respectively were detected by autoradiography. Thus the ~45 kD and the 190 kD polypeptides detected after protein denaturation and immunoprecipitation with the crude anti-Uss1 antiserum were artefactually isolated with this antiserum. The 45 kD polypeptide could have been Prp24p remaining associated with the denatured Uss1p since both proteins are present in the U6 snRNP particle co-immunoprecipitated with anti-Prp4 or with anti-Uss1 antiserum.

The strongest protein-U6 RNA UV-crosslinking signals were obtained when the U6 RNA was radiolabelled with $^{32}$p-αUTP, while much weaker signals were detected when

the U6 RNA was radiolabelled with $^{32}$p-αATP or with $^{32}$p-αCTP. This difference was particularly striking for the 10 kD signal which was visible only when U6 RNA ($^{32}$p-αUTP) was used. The relatively inefficient UV-crosslinking of proteins to the U6 RNA radiolabelled with $^{32}$p-αATP confirmed that purines were not the most photoreactive bases. Although cytidine was theoretically much more reactive than adenosine, proteins were as inefficiently UV-crosslinked to the U6 RNA ($^{32}$p-αCTP) as to the U6 RNA ($^{32}$p-αATP). The best UV-crosslinking results were obtained with the U6 ($^{32}$p-αUTP) indicating that, in addition to being a photoreactive base, proteins were contacting the U6 RNA preferentially at uridine residues rather than at cytidine. In particular the 10 kD polypeptide was more likely to bind the U6 RNA in a U-rich sequence.

The study of the proteins UV-crosslinked to the U6 RNA confirmed that Prp24p was in close proximity to the U6 RNA in the U6 snRNP (Jandrositz and Guthrie, 1995; Ghetti *et al.,* 1995; this study). Although we failed to confirm the identity of the 45 kD band UV-crosslinked to the U6 RNA and co-immunoprecipitated with anti-Uss1 antibodies, as this antiserum co-immunoprecipitate the U6 snRNP and the U4/U6 snRNP particles, it is highly probable that this protein was actually Prp24p.

Uss1p is an Sm-like U6 snRNP specific protein which shares significant homology with the yeast SmD3 core protein (Cooper *et al.*, 1995). To date, the only known interaction between a core protein and snRNA is that between the 5' moiety of the U1 snRNA Sm site and the G protein. This contact was demonstrated by UV-crosslinking experiments in HeLa cells (Heinrichs *et al.*, 1992). Uss1p does not contain a good match to any of the characterised RNA binding motifs, making it hard to predict its direct binding with U6 snRNA. The UV-crosslinking experiments described in this work indicate that actually Uss1p contacts the U6 RNA in the U6 snRNP, U4/U6 snRNP and in the tri-snRNP. Uss1p is required for U6 snRNA stability (Cooper *et al.*, 1995); it may stabilises the U6 RNA molecule through its direct binding.

Prp8p is exceptionally large (280 kD) and its size and sequence are highly conserved (Hodges *et al*, 1995), suggesting it has an important role to play. Although Prp8p

contains no obvious similarities to known RNA-binding motifs, it is known to bind at least three different RNAs: exon1 of the pre-mRNA; the 3' splice site of the lariat intermediate (Teigelkamp *et al.*, 1995, 1995b); U5 snRNA (Ian Dix *et al.*, 1998). Now a fourth interaction has been demonstrated with U6 RNA; indeed we have shown that Prp8p contacts the U6 RNA in the U4/U6.U5 tri-snRNP. Prp8p may stabilise U6 snRNA-RNA interaction(s) as has been proposed for the U5 snRNA and the 5' and 3' splice site contacts in the spliceosome (Beggs *et al.*, 1995; Teigelkamp *et al.*, 1995).

In addition to Prp24p, Uss1p and Prp8p, other unidentified polypeptides were in close proximity to the U6 RNA (see Table 12). Some of them contacted the U6 RNA in all snRNP particles, while others seem to release from the U6 RNA (or to be more loosely associated with the U6 RNA) as snRNP particle assembly progressed and other proteins seem to contact the U6 RNA only in the tri-snRNP. From these observations and by assuming that proteins of similar size and precipitated with different antisera correspond to identical entities, we can speculate that a pool of protein (at least 10 kD, 16 kD, Uss1p, 70-75 kD) contacts the U6 RNA in the U6 snRNP and remain associated at least until in the tri-snRNP particle while other proteins move away from or contact the U6 snRNA as snRNP assembly progresses. This gives us a first dynamic picture of the U6 RNA-protein interactions detected by UV-crosslinking. Therefore it would be very interesting to follow these contacts in the spliceosome and also to identify the unknown polypeptides detected so far.

# CHAPTER SIX

## Investigation of the Direct Binding of the Sm-like Proteins to the U6 SnRNA

## INTRODUCTION

Prp24p, Uss1p and SmX4p/Uss2p are associated with the U6 snRNA (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995; Ghetti *et al.*, 1995; Cooper *et al.*, 1995; Séraphin, 1995). Five ORFs (YJL124c, YBL026w, YER146w, YDR378c and YNL147w) encoding putative U6 snRNP proteins which share homology with the core proteins have been identified (Séraphin, 1995; Fromont-Racine *et al.*, 1997, see section 1.4.3.2). Two additional ORFs (YJR022c and SMX1/YCR020Ca) encoding yeast proteins containing the Sm motifs but not much other similarity with the core proteins have been identified by computer sequence analysis (Séraphin, 1995; Fromont-Racine *et al.*, 1997). An exhaustive two-hybrid screen showed that the protein encoded by *YJR022w* contacts Uss1p and the putative U6 snRNP-specific Uss5, Uss6, Uss7 proteins (Fromont-Racine *et al.*, 1997; A. Mayes, 1998). Further screens using the Sm-like U6 snRNP proteins as baits have generated an entire network of protein-protein interactions suggesting that all these polypeptides are interacting (directly or indirectly) with each other (A. Mayes, 1998).

The deletion in haploid yeast cells of the ORFs encoding the Sm-like proteins is associated with a lethal or a temperature-sensitive phenotype, correlated (for most of them) to a detrimental effect on the level of U6 snRNA in the cells (A. Mayes, 1998). Immunoprecipitation and Northern blotting analysis showed that the Sm-like proteins are associated with the U6 snRNA. Taken together these data suggest that these polypeptides constitute a set of U6 snRNP specific proteins.

The direct binding of the Sm-like proteins to the U6 snRNA was investigated by UV-crosslinking and the results are presented in this chapter.

A description of the yeast strains used in this chapter is reported in Table 13.

**Table 13. Description of the yeast strains used in this chapter.**

These strains have been constructed by A. Mayes (this Lab; 1998). They are all haploids and contain a deleted ORF encoding an Sm-like U6 snRNP-associated protein. The missing protein is supplied in *trans* by a corresponding fusion protein. The sequence encoding the fusion polypeptide is provided by a plasmid contained in the yeast strain (Table 7). Only the AEMY28 strain does not harbour a deleted ORF. Instead, the sequence encoding the HA-epitope has been inserted onto the yeast genome upstream of the YJL124c ORF.

| Name | ORF deleted | Plasmid complementing the ORF deletion | Particular characteristics |
|---|---|---|---|
| **AEMY42** | YNL147w | pAEM59 | Overproduce LexA-Uss5p (34 kD) |
| **AEMY43** | YNL147w | pAEM62 | Overproduce HA-Uss5p (13 kD) |
| **AEMY44** | YDR378c | pAEM34 | Overproduce LexA-Uss4p (36 kD) |
| **AEMY45** | YDR378c | pAEM61 | Overproduce HA-Uss4p (17 kD) |
| **AEMY31** | YLR438c | pAEM64 | Overproduce HA-Uss2p (20 kD) |
| **AEMY29** | YER146w | pAEM70 | Overproduce Gal4-HA-Uss7p (35 kD) |
| **AEMY28** | No ORF deletion | Insertion of the sequence coding for the HA epitope upstream the YJL124c ORF | Produce HA-Spb8p (22-23 kD) |
| **AEMY46** | YBL026w | pAEM68 | Overproduce Uss6-HAp (13 kD) |
| **LMA4-2A** | YJR022w | pAEM71 | Overproduce Gal4-HA-Uss3p (38 kD) |

## 6.1 Investigation of U6 RNA-Uss5p Contacts

The deletion of the YNL147c ORF in haploid cells is associated with a temperature-sensitive phenotype at 37°C, and a diminution of the level of U6 snRNA is observed at the restrictive temperature. Immunoprecipitation and Northern blotting analysis showed that Uss5p is associated with the U6 snRNA and with the U4 snRNA probably indirectly via its interaction with the U6 snRNA (A. Mayes, 1998).

Uss5p is a 12 kD protein. Its direct binding with the U6 RNA was investigated in two ways (Figure 27).

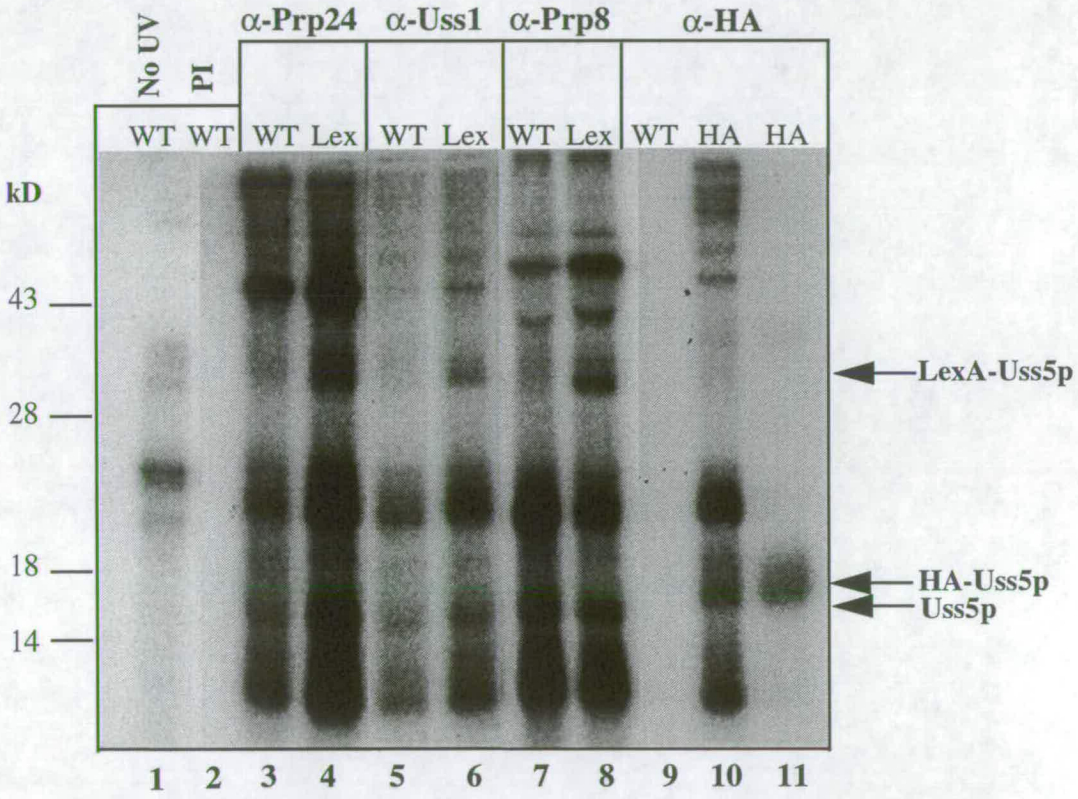
1- Splicing extracts were prepared from AEMY42 yeast cells (Table 13) overproducing a LexA-Uss5 fusion protein. The LexA is located at the N-terminus of Uss5p and the fusion protein possesses a theoretical molecular weight of 34 kD, permitting its identification by size, compared to results with untagged protein.

2- Splicing extracts were prepared from AEMY43 cells (Table 13) overproducing an HA epitope-tagged Uss5 protein. The HA epitope is located at the N-terminus of Uss5p and the fusion protein possesses a calculated molecular weight of ~13 kD (A. Mayes, 1998). This permits its identification by immunoprecipitation with anti-HA antibodies rather than by the very small size difference.

U6 snRNPs were reconstituted with U6 RNA ($^{32}$P-αUTP) in splicing extract containing LexA-Uss5p or in an extract prepared from wild-type SC261 yeast cells. After UV-crosslinking, proteins were immunoprecipitated under nondenaturing conditions with anti-Prp24 (Figure 27; lanes 3 and 4) or anti-Uss1 antibodies (Figure 27; lanes 5 and 6) or with anti-Prp8 antiserum (Figure 27; lanes 7 and 8). The ~16 kD signal detected in a wild-type extract was much weaker in an extract containing the LexA-Uss5 fusion protein, also, an additional band at 30-40 kD was detected (compare lane 3 versus 4, lane 5 versus 6, lane 7 versus 8). This band shift was observed with the protein complexes co-immunoprecipitated with either anti-Prp24, anti-Uss1 or with anti-Prp8 antibodies. Thus LexA-Uss5p contacts the U6 RNA in the U6 snRNP, in the tri-snRNP and probably in the U4/U6 snRNP. U6 snRNPs were reconstituted in a splicing extract

**Figure 27. Analysis of U6 RNA-Uss5p contacts.**

U6 snRNPs were reconstituted (50 µl) with a randomly radiolabelled U6 RNA ($^{32}$P-αUTP) (5 nM) in splicing extract prepared from: SC261 cells; AEMY42 cells overproducing the LexA-Uss5 fusion protein; AEMY43 cells overproducing the HA-Uss5 fusion protein. After UV-crosslinking, proteins were immunoprecipitated under nondenaturing conditions with the affinity purified anti-Prp24 antibodies (lane 1, 3, 4); anti-Uss1 antibodies (lanes 5 and 6); with anti-Prp8 antiserum (lanes 7 and 8) or with monoclonal anti-HA antibodies (lanes 9 and 10). Proteins were denatured then immunoprecipitated with anti-HA antibodies (lane 11). As controls, proteins were incubated with the preimmune serum in sample 2 and UV-crosslinking was omitted in sample 1. Protein complexes were washed (150 mM NaCl) then resolved on a 15% SDS-polyacrylamide gel. WT: proteins extracted from SC261 cells; Lex: extracts prepared from cells overproducing the LexA-Uss5 fusion protein; HA: extracts prepared from cells overproducing the HA-Uss5 fusion protein.


No UV
PI
α-Prp24
α-Uss1
α-Prp8
α-HA
WT WT WT Lex WT Lex WT Lex WT HA HA
kD
43
28
18
14
LexA-Uss5p
HA-Uss5p
Uss5p
1 2 3 4 5 6 7 8 9 10 11

containing the HA-Uss5 fusion protein. After UV-crosslinking, proteins were immunoprecipitated with anti-HA antibodies under nondenaturing conditions (Figure 27; lane 10). At least five polypeptides (10 kD, 17 kD, 25 kD, ~40 kD and 50-60 kD) contacted the U6 RNA in a protein complex associated with HA-Uss5p. Only one protein (17 kD) was detected when the polypeptides were denatured before immunoprecipitation (lane 11), indicating that HA-Uss5p was also in close proximity to the U6 RNA. This fusion protein appears to contact the U6 RNA in a protein complex very similar (with respect to their molecular weight) to those co-immunoprecipitated with anti-Prp24, anti-Uss1 and anti-Prp8 antisera.

Splicing extracts prepared from cells overproducing either LexA-Uss5p or HA-Uss5p were functional before and after U6 snRNP reconstitution (data not shown). Thus the formation of snRNPs and spliceosomes was not affected by the overproduction of these fusion proteins. In extract containing LexA-Uss5p, apart from the 30-40 kD signal reflecting the UV-crosslinking of LexA-Uss5p to the U6 RNA, the other U6 RNA-crosslinked proteins detected with anti-Prp24, anti-Uss1 and anti-Prp8 antibodies were identical (with respect to their size) to the proteins immunoprecipitated with the same antibodies in extract prepared from wild-type cells. It is therefore unlikely that the binding of LexA-Uss5p to the U6 RNA is an artefact of the overproduction of this protein in the cells. Both LexA-Uss5 and the HA-Uss5 fusion proteins bind to the U6 RNA, this contact is unlikely to be an artefact of the extra amino acids added at the N-terminus of Uss5p. These data suggest that the Uss5 protein is in close proximity to the U6 RNA in the U6 snRNP, the U4/U6.U5 tri-snRNP and maybe in the U4/U6 snRNP. The persistence of a fraction of the 16 kD signal in extracts overproducing the LexA-Uss5 protein (and not in extracts producing the untagged Uss5p) suggests that in addition to Uss5p, other protein(s) that co-migrate with Uss5p on a 15% SDS-polyacrylamide gel may contact the U6 RNA.

## 6.2 Investigation of U6 RNA-Uss4p Interactions

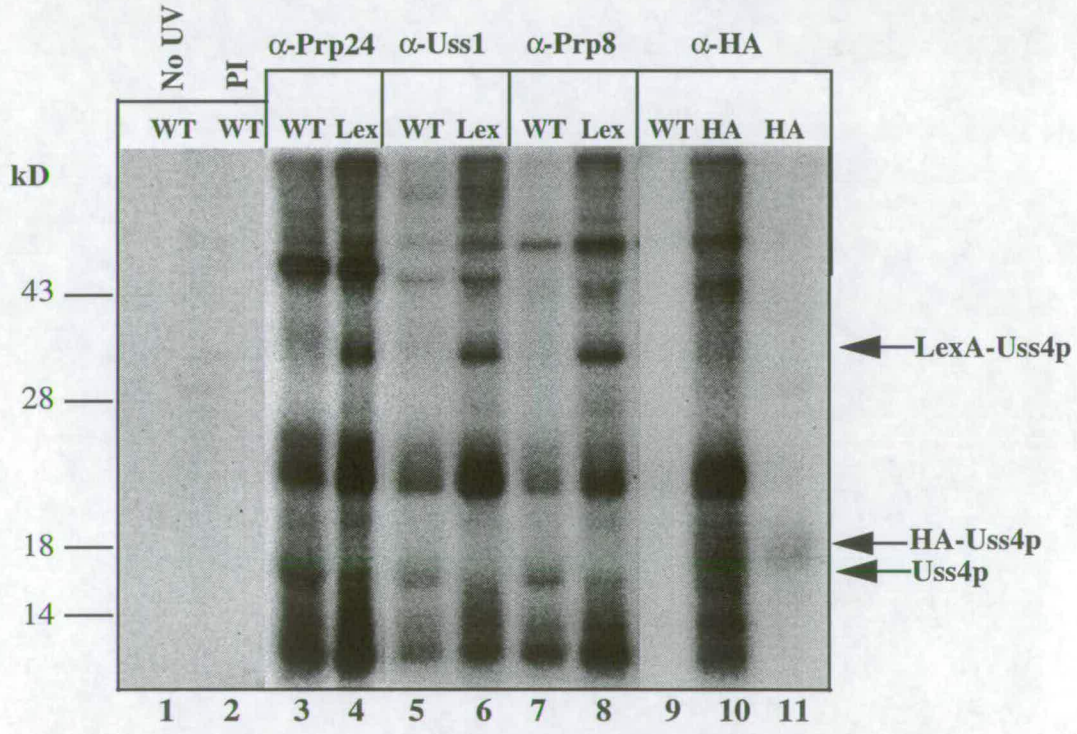
YDR378c encodes a 14 kD protein which shares homology with the core SmF protein (Fromont-Racine *et al.*, 1997). The deletion of the YDR378c ORF in haploid cells is associated with a temperature-sensitive phenotype at 37°C correlated with a decrease in the stability of the U6 snRNA. Immunoprecipitation and Northern blotting analysis showed that Uss4p is associated with the U6 snRNA and with the U4 snRNA, probably through its interaction with the U6 snRNA (A. Mayes, 1998).

Two splicing extracts were prepared. One was made from AEMY44 (Table 13) cells overproducing a LexA-Uss4 fusion protein which possesses a molecular weight of 36 kD. The second extract was made from AEMY45 (Table 13) cells which overproduce an HA epitope-tagged Uss4 fusion protein of 17 kD.

The U6 RNA-Uss4p contacts were investigated by UV-crosslinking (Figure 28). The detected signals are UV and antisera specific (lanes 1, 2). In extracts overproducing the LexA-Uss4 fusion protein, a signal at ~35 kD was noticed while a very weak signal remained at ~16 kD. This 35 kD protein was not detected when U6 snRNPs were reconstituted in an extract prepared from the wild-type SC261 cells (compare lanes 3 and 4, 5 and 6, 7 and 8). This difference was observed in the protein complexes co-immunoprecipitated with the three varieties of antibodies tested, indicating that LexA-Uss4p contacts the U6 RNA in the U6 snRNP, the U4/U6.U5 tri-snRNP and maybe in the U4/U6 snRNP. In extracts prepared from cells overproducing HA-Uss4p, several proteins (10 kD, 18 kD, 25 kD, ~45 kD, 60 kD and other signals of undefined higher molecular weights) were co-immunoprecipitated with anti-HA antibodies (Figure 28, lane 10). Only the ~18 kD signal remained after immunoprecipitation of denatured proteins (lane 11). Thus HA-Uss4p is in close proximity to the U6 RNA in a complex of proteins very similar (with respect to their apparent molecular weight) to those precipitated with anti-Prp24, anti-Uss1 or with the anti-Prp8 antibodies.

**Figure 28. Analysis of U6 RNA-Uss4p contacts.**

U6 snRNPs were reconstituted (50 µl) with 5 nM of randomly radiolabelled U6 RNA ($^{32}$P-αUTP) in a splicing extract prepared from: SC261 cells; AEMY44 cells overproducing the LexA-Uss4 fusion protein; AEMY45 cells overproducing the HA-Uss4 fusion protein. After UV-crosslinking, proteins were immunoprecipitated under nondenaturing conditions with the affinity-purified anti-Prp24 (lane 1, 3, 4) or anti-Uss1 antibodies (lanes 5 and 6) or with anti-Prp8 antiserum (lanes 7 and 8) or with monoclonal anti-HA antibodies (lanes 9 and 10). Proteins were denatured then immunoprecipitated with anti-HA antibodies (lane 11). As controls, proteins were incubated with preimmune serum in sample 2 and UV-crosslinking was omitted in sample 1. Protein complexes were washed (150 mM NaCl) then resolved on a 15% SDS-polyacrylamide gel. WT: proteins extracted from SC261 cells; Lex: extracts prepared from cells overproducing the LexA-Uss4 fusion protein; HA: extracts prepared from cells overproducing the HA-Uss4 fusion protein.


No UV
PI
α-Prp24
α-Uss1
α-Prp8
α-HA
WT
WT
WT
Lex
WT
Lex
WT
Lex
WT
HA
HA
kD
43
28
18
14
LexA-Uss4p
HA-Uss4p
Uss4p
1
2
3
4
5
6
7
8
9
10
11

These data demonstrate that LexA-Uss4p and HA-Uss4p bind to the U6 RNA. For the same reasons as were indicated for the binding of Uss5p to the U6 RNA, the interaction of the LexA-Uss4p and the HA-Uss4p fusion proteins with the U6 RNA probably indicates that Uss4p itself contacts the U6 RNA.

## 6.3 Analysis of U6 RNA Contacts with Spb8p, Uss6p, Uss2p, Uss7p and Uss3p

The deletion of the YJL124c ORF in cells is associated with a temperature-sensitive phenotype at 37°C which is not linked with the instability of any spliceosomal snRNAs. Immunoprecipitation and Northern blotting analysis showed that the encoded Spb8 protein is only weakly associated with the U6 snRNA (not above 50 mM NaCl) (A. Mayes, 1998).

The deletions of YBL026w, YLR438c and YER146w ORFs are associated with a lethal phenotype. Temperature-sensitive or galactose-inducible alleles of these genes have been generated (A. Mayes, this lab), and an arrest of growth was observed in cells shifted to the nonpermissive temperature or to glucose-containing medium. This effect was correlated with a decrease in the levels of U6 and $U5_L$ snRNAs in AEMY30 (YBL026wΔ::*HIS3*; pAEM55) cells (Tables 7 and 9) or in AEMY31 (YLR438cΔ::*HIS3*; pAEM64) cells grown respectively at the nonpermissive temperature or in glucose-containing medium (A. Mayes, 1998). The level of the spliceosomal snRNAs remained unchanged in AEMY29 (YER146wΔ::*TRP1*; pAEM70) cells grown at nonpermissive temperature. Immunoprecipitation and Northern blotting analysis showed that all three proteins were associated with the U6 snRNA and with the U4 snRNA, probably through their interaction with the U6 snRNA. Uss6p was also associated with a tiny amount of U5 snRNA (A. Mayes, 1998).

Deletion of the YJR022w ORF was lethal and no conditional allele was available. Immunoprecipitation data indicate that the corresponding protein is associated with the U6 snRNA and the U4 snRNA, probably through its association with the former.

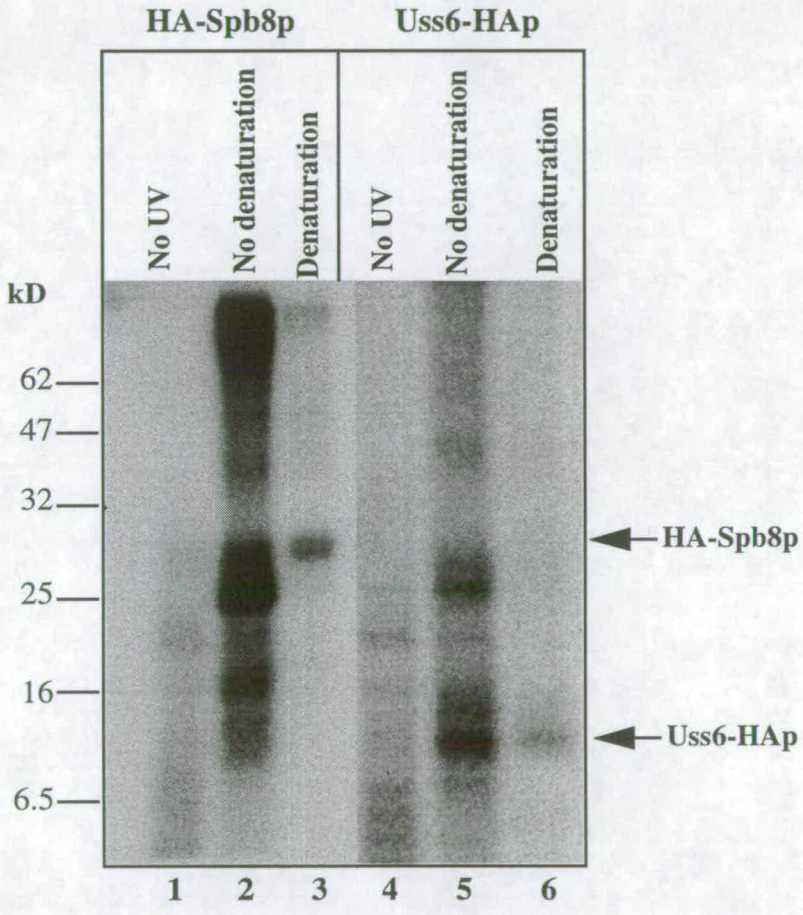
### 6.3.1 Investigation of U6 RNA contacts with Spb8p and Uss6p

AEMY28 cells (Table 14) produce an HA-tagged Spb8 protein with molecular weight 22-23 kD (A. Mayes, 1998). AEMY46 (Table 13) cells grown overproduce an HA-tagged Uss6 protein of a size (~13 kD) (A. Mayes, 1998). As little size difference exists between the wild-type and the fusion proteins, in splicing extracts prepared from either AEMY28 or AEMY46 cells, HA-Sbp8p or Uss6-HAp will be identified by immunoprecipitation with anti-HA antibodies only.

U6 snRNPs were reconstituted with U6 RNA ($^{32}$P-αUTP) in splicing extracts prepared from AEMY28. Subsequently proteins were UV-crosslinked to the nucleic acids and immunoprecipitated with anti-HA antibodies under nondenaturing conditions or after protein denaturation (Figure 29, lanes 1-3). In an extract producing the HA-Spb8 fusion protein, two bands (~16 kD and a large signal at 25-27 kD) were detected after UV-crosslinking (lane 2). Additional signal(s) of undetermined size(s) were visible at the top of the SDS-polyacrylamide gel (lane 2). Only a band at ~27 kD was detected when denatured proteins were immunoprecipitated with anti-HA antibodies (lane 3). HA-Spb8p contacts the U6 RNA in a protein complex quite different (with respect to their molecular weights) than those co-immunoprecipitated with anti-Prp24, anti-Uss1 or with anti-Prp8 antibodies. The expression of *HA-YJL124c* is under the control of the endogenous promoter, therefore as much HA-Spb8p is produced in AEMY28 cells as Spb8p is produced in wild-type yeast cells. The binding of HA-Spb8p cannot be an artefact due to overproduction. The interaction of HA-Uss5p and HA-Uss4p with the U6 RNA was not linked to the presence of the HA epitope, thus it is unlikely that the HA epitope has promoted the binding of HA-Spb8p to the U6 RNA. The UV-

**Figure 29. Analysis of U6 RNA-Spb8p and -Uss6p contacts.**

U6 snRNPs were reconstituted (50 µl) with a randomly radiolabelled U6 RNA ($^{32}$P-αUTP) (5 nM) in a splicing extract prepared from AEMY28 cells producing an HA-Spb8 fusion protein (lanes 1, 2, 3); or from AEMY46 cells overproducing the Uss6-HA fusion protein (lanes 4, 5, 6). After UV-crosslinking, proteins were immunoprecipitated under nondenaturing conditions with anti-HA antibodies (lanes 2 and 5), or proteins were denatured before immunoprecipitation with anti-HA antibodies (lanes 3 and 6). Samples were not UV-irradiated in lanes 1 and 4. Protein complexes were washed (150 mM NaCl) then resolved on a 15% SDS-polyacrylamide gel.


HA-Spb8p
Uss6-HAp
No UV
No denaturation
Denaturation
No UV
No denaturation
Denaturation
kD
62
47
32
25
16
6.5
HA-Spb8p
Uss6-HAp
1
2
3
4
5
6

crosslinking data presented here suggest that HA-Spb8p and probably also the Spb8 protein contact directly the U6 RNA.

U6 snRNPs were reconstituted in an extract prepared from cells overproducing the Uss6-HA protein. After UV-crosslinking, proteins were immunoprecipitated with anti-HA antibodies (Figure 29, lanes 4-6). Radioactive signals were detected in the absence of UV-irradiation (lane 4). They probably correspond to fragments of undigested U6 RNA interacting with proteins cross-reacting with the anti-HA antibodies, or these RNA fragments may interact with the PAS beads despite the washes (150 mM NaCl). Few proteins contacting the U6 RNA were associated with the Uss6-HA fusion protein (lane 5); only two signals (14 kD and 26 kD) were observed. The 14 kD band was detected when denatured proteins were immunoprecipitated with anti-HA antibodies (lane 6), suggesting that Uss6-HAp binds, although weakly, to the U6 RNA.

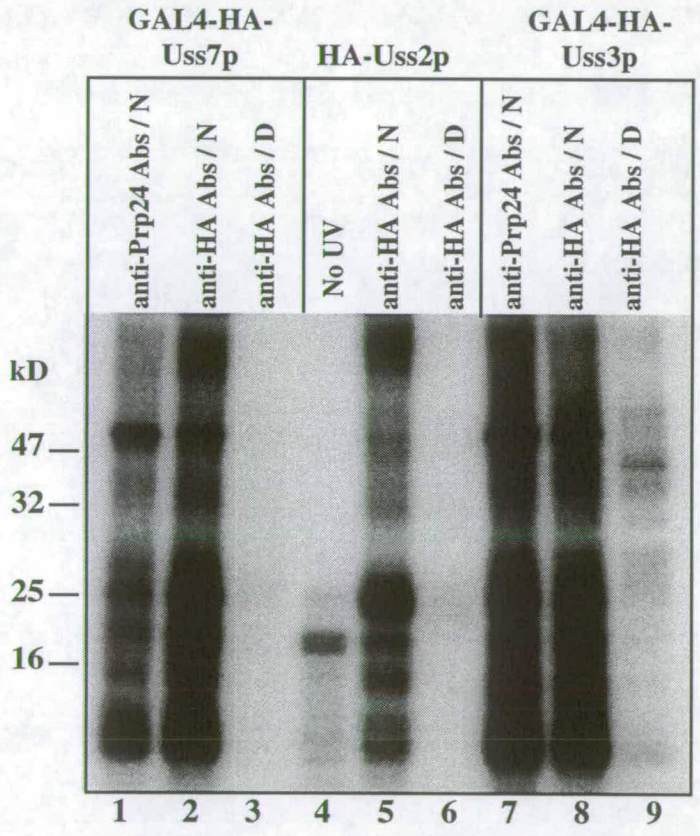
### 6.3.2 Investigation of the U6 RNA-Uss7p, -Uss2p and -Uss3p contacts

Splicing extracts were prepared from AEMY31 (Table 13) cells overproducing an HA epitope-tagged Uss2 protein of 20 kD. AEMY29 (Table 13) cells overproducing a Gal4-HA-Uss7 fusion protein of 35 kD. LMA4-2A (Table 13) cells overproducing a Gal4-HA-Uss3 fusion protein of 38 kD. Since there is only a small size difference between HA-Uss2p and Uss2p, the former can only be identified using anti-HA antibodies. The greater size difference (15 kD-22 kD) between Uss7p or Uss3p and their respective Gal4-HA-fusion proteins allowed their identification by size, since no polypeptide of 35 kD or 38 kD contacts the U6 RNA in protein complexes isolated with anti-Uss1, anti-Prp24 or anti-Prp8 antibodies (see Figure 25). These fusion proteins can also be identified using anti-HA antibodies.

In extracts containing either Gal4-HA-Uss7p or HA-Uss2p, after U6 snRNP reconstitution and UV-crosslinking, no denatured proteins were immunoprecipitated with the anti-HA antibodies (Figure 30; lanes 3 and 6). No signal the size of the Gal4-

**Figure 30. Analysis of U6 RNA-Uss7p, -Uss2p and -Uss3p contacts.**

U6 snRNPs were reconstituted (50 µl) with a randomly radiolabelled U6 RNA ($^{32}$P-αUTP) (5 nM) in a splicing extract prepared from AEMY29 cells producing an Gal4-HA-Uss7 fusion protein (lanes 1-3); from AEMY31 cells overproducing the HA-Uss2 fusion protein (lanes 4-6); from LMA4-2A cells overproducing the Gal4-HA-Uss3 fusion protein (lanes 7-9). After UV-crosslinking, proteins were immunoprecipitated under nondenaturing conditions with anti-HA (lanes 2, 5 and 8) or with anti-Prp24 antibodies (lanes 1 and 7). Proteins were denatured then immunoprecipitated with anti-HA antibodies (lanes 3, 6 and 9). Samples were not UV-irradiated in lane 4. Protein complexes were washed (150 mM NaCl) then resolved on a 15% SDS-polyacrylamide gel.


GAL4-HA-
Uss7p
HA-Uss2p
GAL4-HA-
Uss3p
anti-Prp24 Abs / N
anti-HA Abs / N
anti-HA Abs / D
No UV
anti-HA Abs / N
anti-HA Abs / D
anti-Prp24 Abs / N
anti-HA Abs / N
anti-HA Abs / D
kD
47
32
25
16
1
2
3
4
5
6
7
8
9

HA-Uss7 protein (35 kD) was detected in the protein complexes isolated with anti-Prp24 (lane 1) or with anti-HA antibodies (lane 2). The 10 kD band was not affected by the replacement in the cells of Uss2p or Uss7p by their respectively fusion proteins of a higher molecular weight. Thus under the experimental conditions tested, neither HA-Uss2p nor Gal4-HA-Uss7p bind to the U6 RNA.

In extract overproducing the Gal4-HA-Uss3 protein, after U6 snRNP reconstitution and UV-crosslinking, a very faint signal the size of the fusion protein (38 kD) was detected when proteins were co-immunoprecipitated with anti-Prp24 or with anti-anti-HA antibodies (Figure 30, lanes 7 and 8). An overexposure of the autorad was necessary to visualize this signal. A very weak signal the size of Gal4-HA-Uss3p was detected when after UV-crosslinking denatured proteins were immunoprecipitated with anti-HA antibodies (lane 9). Therefore presumably, Gal4-HA-Uss3p and probably also Uss3p contacts the U6 RNA but crosslinks very weakly.

## 6.4 Investigation of the Binding Specificity of the Proteins to the U6 RNA

### 6.4.1 The binding specificity of Spb8p

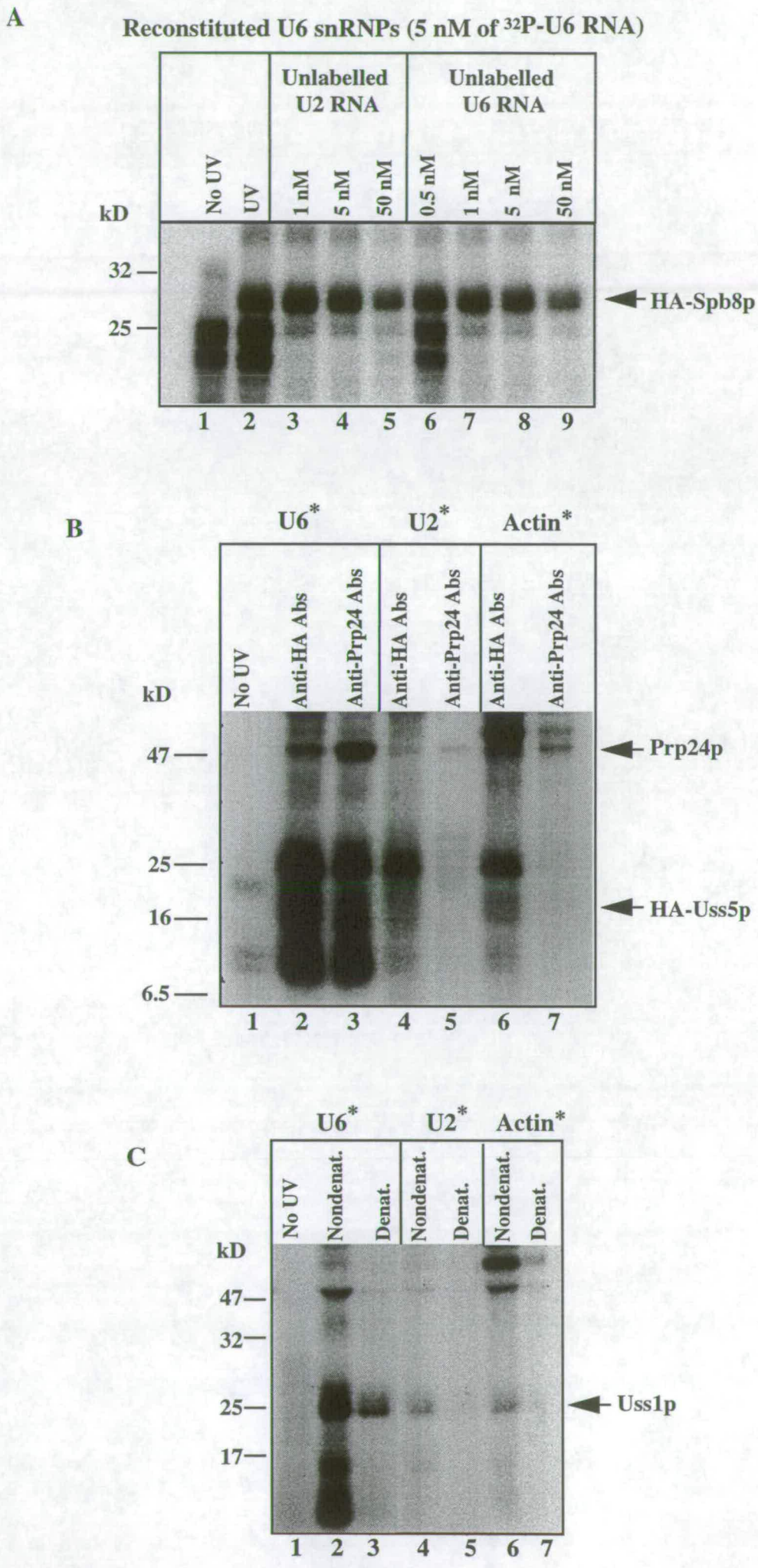
In a splicing extract prepared from AEMY28 cells which produce the HA-Spb8 fusion protein, U6 snRNPs were reconstituted with 5 nM of U6 RNA ($^{32}$P-αUTP). Increasing amounts of unlabelled U6 or U2 RNAs were added in the reconstitution reaction as competitor RNAs. After UV-crosslinking, the proteins were denatured and immunoprecipitated with anti-HA antibodies (Figure 31, Panel A). Radioactive signals (between 18 kD and 25 kD) were detected in the absence of UV-irradiation, and also in UV-crosslinked samples that contain no or very little (0.5 nM) competitor RNAs (lanes 2 and 6). These signals were probably fragments of radioactive U6 RNA which strongly associate with proteins cross-reacting with the anti-HA antibodies. Alternatively, these RNA fragments may have been attached to the PAS beads despite the washes. The addition of unlabelled U2 or U6 RNAs (1 nM-50 nM) competes for their binding and

**Figure 31. Analysis of the binding specificity of the proteins to the U6 RNA.**

**A:** In an extract prepared from AEMY28 cells producing the HA-Spb8 fusion protein, U6 snRNPs were reconstituted with 5 nM of U6 RNA ($^{32}$P-αUTP) (lanes 1-9). Different amounts of unlabelled U2 RNA (lanes 3-5) or U6 RNA (lanes 6-9) were added at the same time in the reconstitution reaction (50 µl). Samples were UV-irradiated, proteins denatured and immunoprecipitated with anti-HA antibodies. Sample 1 was not submitted to UV-crosslinking.

**B:** In splicing extracts prepared from AEMY43 cells overproducing the HA-Uss5 fusion protein, the endogenous U6 snRNA was depleted by RNase H degradation. 5nM of radiolabelled ($^{32}$P-αUTP) U6 RNA (lanes 1-3); or 5 nM of radiolabelled U2 RNA (lanes 4, 5) was added. 5 nM of radiolabelled truncated actin RNA was added in an extract lacking the endogenous U6 snRNA (lanes 6, 7). Samples were UV-irradiated (except for sample 1), then proteins were immunoprecipitated under nondenaturing conditions with anti-HA antibodies (lanes 1, 2, 4, 6) or with the affinity-purified anti-Prp24 (lanes 3, 5, 7) antibodies. Protein complexes were washed with 150 mM NaCl and resolved on a 15% SDS-PAGE gel.

**C:** The endogenous U6 snRNA was depleted in a splicing extract prepared from SC261 wild-type cells. Then 5nM of radiolabelled ($^{32}$P-αUTP) U6 RNA (lanes 1-3); 5nM of radiolabelled U2 RNA (lanes 4, 5); 5nM of radiolabelled actin RNA harboring Exon1 and a short intron sequence finishing after the branchpoint (lanes 6, 7) was added. Samples were UV-irradiated (except for sample 1), then proteins were immunoprecipitated under nondenaturing conditions or after proteins denaturation with the affinity-purifed anti-Uss1 antibodies. Protein complexes were washed with 150 mM NaCl and resolved on a 15% SDS-PAGE gel.


A
Reconstituted U6 snRNPs (5 nM of 32P-U6 RNA)
Unlabelled U2 RNA
Unlabelled U6 RNA
No UV
UV
1 nM
5 nM
50 nM
0.5 nM
1 nM
5 nM
50 nM
kD
32
25
HA-Spb8p
1 2 3 4 5 6 7 8 9
B
U6*
U2*
Actin*
No UV
Anti-HA Abs
Anti-Prp24 Abs
Anti-HA Abs
Anti-Prp24 Abs
Anti-HA Abs
Anti-Prp24 Abs
kD
47
25
16
6.5
Prp24p
HA-Uss5p
1 2 3 4 5 6 7
C
U6*
U2*
Actin*
No UV
Nondenat.
Denat.
Nondenat.
Denat.
Nondenat.
Denat.
kD
47
32
25
17
Uss1p
1 2 3 4 5 6 7

no RNA fragments were detected any more (Figure 31A; lanes 3-5 and lanes 7-9). The addition of increasing amounts of unlabelled U6 RNA (0.5 nM-50 nM) in the U6 snRNP reconstitution reaction, affected slightly the binding of Spb8p to the radiolabelled U6 RNA (lanes 6-9). Only a three fold reduction of the UV-crosslinking signal was observed when a ten fold excess of unlabelled U6 RNA was present in the reaction (lane 9). This indicates that the level of reconstituted U6 snRNPs was not saturated in the presence of 5 nM of U6 RNA ($^{32}$P-αUTP). Therefore, the addition of unlabelled U6 RNA initially increases the number of reconstituted U6 snRNPs, prior to competing with the radiolabelled U6 RNA for its association with proteins. This result was predictable since it was shown (Figure 19) that only ~60 % of the initial splicing activity of the extract was recovered following U6 snRNP reconstitution with 5 nM of U6 RNA. Larger amounts of U6 RNA (50 nM-100 nM) were required for a complete recovery of splicing. In this experiment the amount of reconstituted U6 snRNPs was limited, and therefore the ability of the unlabelled U6 RNA to compete for the binding of Spb8p with the radiolabelled U6 RNA cannot be investigated. Consequently no comparison with the binding of Spb8p with the U2 snRNA in the absence of the endogenous U6 snRNA can be made. These experiments were inconclusive.

### 6.4.2 The binding specificity of Uss5p and Prp24p

A splicing extract was prepared from AEMY43 cells which overproduce an HA-Uss5 fusion protein of 18 kD. In this extract, U6 snRNP particles were reconstituted either with a radiolabelled U6 RNA or with a radiolabelled U2 RNA. U6 snRNPs were also reconstituted with a nonfunctional radiolabelled actin substrate (produced by run-off transcription of p283 linearised with *Cla*I, Table 9). This actin RNA is truncated and possesses exon1 and a short intron sequence finishing just after the branchpoint. Subsequently, proteins were UV-crosslinked to the RNAs and immunoprecipitated with anti-HA or anti-Prp24 antibodies under nondenaturing conditions (Figure 31, Panel B). When the radiolabelled U6 RNA was added, HA-Uss5p as well as three other polypeptides (10 kD, 25 kD, ~50 kD), were co-immunoprecipitated with anti-HA

antibodies (lane 2). In addition to Prp24p, other polypeptides (10 kD, 25 kD, a weak 16 kD band and a rather diffuse signal at 18 kD) were UV-crosslinked to the U6 RNA and co-immunoprecipitated with anti-Prp24 antibodies (lane 3). When instead a radiolabelled U2 RNA was added in the reaction, a 25 kD protein and a weak 50 kD signal were detected after immunoprecipitation with anti-HA antibodies (lane 4). Only a faint 50 kD signal was detected after immunoprecipitation with anti-Prp24 antibodies (lane 5). When a truncated radiolabelled actin RNA was added in an extract lacking the endogenous U6 snRNA, two bands (25 kD and >50 kD) and a rather diffuse 18 kD signal were detected after UV-irradiation and immunoprecipitation with anti-HA antibodies (lane 6). Two proteins (50 kD and >50 kD) were co-immunoprecipitated with anti-Prp24 antibodies after UV-crosslinking (lane 7). The 18 kD signal might reflect a weak binding of HA-Uss5p to the actin RNA or the interaction of a polypeptide of a similar size. Even if HA-Uss5p binds to the actin RNA it shows a much greater preference for the U6 RNA.

Among the proteins (10 kD, 25 kD and 50 kD) associated with HA-Uss5p and UV-crosslinked to the U6 RNA, the 10 kD protein bound neither the U2 RNA nor the actin RNA, while the 25 kD polypeptide contacted all the different RNAs tested. The faint 50 kD signal detected after UV-crosslinking with the U2 or actin RNAs and immunoprecipitated with anti-HA and anti-Prp24 antibodies, suggests that either Prp24p interacts very weakly with these RNAs or a polypeptide of a size similar to Prp24p binds to the U2 and/or actin RNAs. Prp24p possesses three RNA binding domains (Shannon and Guthrie, 1991) and may therefore contact other RNAs in the absence of U6 snRNA. Even if Prp24p contacts the U2 or actin RNAs, it shows a much greater preference for the U6 RNA.

### 6.4.3 Specificity of the binding of Uss1p

Splicing extracts prepared from wild-type SC261 cells (Table 7) were depleted from their endogenous U6 snRNA. Equal amounts of either radiolabelled U6 or U2 RNAs, or a truncated actin RNA were then added. Proteins were UV-crosslinked to the nucleic acids and immunoprecipitated with anti-Uss1 antibodies (Figure 31, Panel C). Uss1p and other proteins (10 kD, 16 kD, 50 kD (this signal probably corresponds to the binding of Prp24p to the U6 RNA, see chapter 5)) were UV-crosslinked to the radiolabelled U6 RNA and co-immunoprecipitated with anti-Uss1 antibodies (lane 2). When a radiolabelled U2 RNA was added in the extract lacking the endogenous U6 snRNA, after UV-crosslinking, two faint signals were detected at 25 kD and at 50 kD (lane 4). The same proteins and a third one of >50 kD were UV-crosslinked to the radiolabelled actin RNA (lane 6). None of these polypeptides was detected when denatured proteins were immunoprecipitated with anti-Uss1 antibodies (lanes 5 and 7). Anti-Uss1 antibodies precipitate the denatured Uss1p less efficiently (compare lanes 2 and 3). Thus we cannot rule out that despite the loss of signal, the 25 kD band UV-crosslinked to the radiolabelled U2 or actin RNAs and co-immunoprecipitated with anti-Uss1 antiserum, was Uss1p very weakly associated with these RNAs. Alternatively, this 25 kD band may reflect the binding of another protein. Even if Uss1p contacted the U2 and actin RNAs, it shows a clear preference for the U6 RNA. If another polypeptide of ~25 kD contacted the U2 and actin RNAs, a potential candidate would be Sbp8p, since it possesses a size (20 kD) similar to Uss1p (21 kD) and does not show much specificity for the binding to the U6 RNA (see Figure 31A). As discussed in the previous section (9.4.2), the 50 kD band might correspond to Prp24p weakly associated with the U2 or actin RNAs or to an other polypeptide.

## 6.5 Discussion

The investigation of U6 RNA-protein interactions with the Sm-like U6 snRNP proteins identified the 16 kD signal. Two proteins: Uss4p (14 kD) and Uss5p (12 kD) contacted the U6 RNA and migrated close together as a ~16 kD crosslinked signal on a 15% SDS-polyacrylamide gel. These polypeptides possessed a molecular weight higher than calculated. The covalent binding of an RNA molecule to these small proteins probably affects their mobility so that Uss4p and Uss5p appear much bigger. Both Uss4p and Uss5p contact the U6 RNA in the U6 snRNP, the U4/U6.U5 tri-snRNP and probably in the U4/U6 snRNP, through their direct contact with the U6 RNA. This result is in good agreement with the association of both HA epitope-tagged proteins with the U6 and the U4 snRNAs in yeast cells (A. Mayes, 1998). No association between Uss4p or Uss5p and the U5 snRNA has been observed in yeast (A. Mayes, 1998). This could reflect a detrimental effect of the HA epitope on the association of these proteins with the tri-snRNP, or the HA epitope may be masked in the tri-snRNP. Uss4p and Uss5p are required for the stability of the U6 snRNA (A. Mayes, 1998). They stabilise this molecule probably through their direct contact.

The 10 kD band was not detected any more when RNase digestion was followed by proteinase K treatment (40 min incubation at 37°C, data not shown). This signal may correspond to a fragment of U6 RNA protected from complete RNase degradation by UV-crosslinked protein(s). The degradation of the proteins by proteinase K would therefore made it susceptible to total RNase digestion. The proteinase K sensitivity of the 10 kD band may also indicate that very small protein(s) were UV-crosslinked to the U6 RNA. Among the U6 snRNP-associated proteins, Uss6p (11 kD), Uss2p (10 kD) and Uss7p (10 kD) could potentially contact the U6 RNA. UV-crosslinking analyses have shown that the HA epitope-tagged Uss6 protein contacts the U6 RNA, but the signal was quite faint and clearly needs to be repeated for confirmation. Western blotting analysis indicated that among all the HA epitope-tagged Sm-like proteins produced in yeast cells, Uss6p was least abundant (A. Mayes, personnal communication). Uss6-HAp was probably unstable and/or the expression of the

encoding gene was inefficient, which correlates with the observation that the expression of *YBL026w-HA* in AEMY46 cells (Table 14) is associated with a temperature-sensitive phenotype at 37°C (A. Mayes, 1998). Thus a detrimental effect of the HA epitope on the function of the protein, or the low abundance of Uss6-HAp in cells, might be responsible for the weak UV-crosslinking of Uss6-HAp to the U6 RNA.

No direct binding between HA-Uss2p and the U6 RNA has been detected under the experimental conditions tested. However, this fusion protein was associated with polypeptides UV-crosslinked to the U6 RNA and co-immunoprecipitated with anti-HA antibodies. The detection of HA-Uss2p (~20 kD) on an SDS-PAGE gel would have been prevented by the presence of other protein(s) (i.e. Uss1p) migrating to a similar position on the gel. If several proteins UV-crosslinked to the U6 RNA co-migrated with Uss2p, in an extract containing only HA-Uss2p, the detection of the loss of the untagged Uss2p signal would have been prevented by the presence of the other protein(s). Anti-HA antibodies precipitate denatured proteins with a low efficiency. Consequently, the failure to detect any denatured protein UV-crosslinked to the U6 RNA using these antibodies, does not absolutely indicate a lack of binding. For all these reasons we cannot declare that HA-Uss2p does not contact the U6 RNA.

No direct binding of Gal4-HA-Uss7p was detected when after UV-crosslinking denatured proteins were immunoprecipitated with anti-HA antibodies. Also no signal the size of the Gal4-HA-Uss7 fusion protein (35 kD) was detected when anti-Prp24 antibodies were used to isolate the UV-crosslinked proteins. This suggests that this fusion protein did not contact the U6 RNA. The overexpression of the Gal4-HA-YER146w ORF complements the lethal phenotype associated with the deletion of YER146w, however, cells show a temperature-sensitive phenotype at 37°C (A. Mayes, 1998). Thus the addition of extra amino acids at the N-terminus of Uss7p might affect its function and maybe also its binding with the U6 RNA.

In extracts overexproducing the Gal4-HA-Uss3 fusion protein a very weak UV-crosslinking signal was detected when denatured proteins were immunoprecipitated

with anti-HA antibodies. Only a faint band the size of Uss3p (38 kD) was detected (after overexposure of the autorad) in protein complexes co-immunoprecipitated with anti-Prp24 or anti-HA antibodies. Thus if the Gal4-HA-Uss3 protein contacted the U6 RNA, this interaction was very weak. This suggests that the presence of additional amino acids at the N-terminus of Uss3p affects its function and/or this protein is loosely associated with the U6 RNA.

The HA-Spb8 protein was UV-crosslinked to the U6 RNA. Although the experiments investigating its binding specificity with the U6 RNA were inconclusive, Spb8p binds to the U2 RNA in the absence of the endogenous U6 snRNA. The size difference (~3 kD) between the wild-type Spb8 and the fusion protein prevented the investigation of its binding in the snRNPs. The expression of *HA-YJL124c* was under the control of the endogenous promoter therefore HA-Spb8p was not overproduced in the cells. Thus the level of production of HA-Spb8p was probably not responsible for its binding to the U2 RNA. Spb8p has been identified by computer analysis as an Sm-like protein (Fromont-Racine *et al.*, 1997). There is only little evidence to suggest that Spb8p is a U6 snRNP-specific protein. From all the Sm-like U6 snRNP proteins tested, Spb8p is the only one showing a weak interaction with the U6 snRNA (it binds at 50 mM NaCl only while the others bind at 150 mM NaCl; A. Mayes, 1998). Also the absence of Spb8p in yeast cells does not affect the stability of the spliceosomal snRNAs (A. Mayes, 1998). Very recently a mutation in the *YJL124c* gene (also named *SPB8)* has been isolated as a suppressor of a poly(A)-binding protein (*PAB1*) gene deletion (Boeck *et al.*, 1998). In the *spb8-2* strain stabilisation of the cap structure is observed, suggesting that Spb8p is involved in the degradation of the mRNAs in the cell (Boeck *et al.*, 1998). Taken together these data suggest that either Spb8p is not a U6 snRNP specific protein, or it possesses a double function and it is also required for the degradation of the mRNAs. This second hypothesis implies that Spb8p is present in the nucleus of the cell to associate with the U6 snRNP, and also in the cytoplasm to contribute to the degradation of mRNAs. Thus, additional factor(s) might be required to either maintain Spb8p in the cytoplasm or for its nuclear import. In a splicing extract, nuclear and cytoplasmic components are mixed together, therefore the lack of membrane separating the nuclear

and the cytoplasmic components, or the loss of factor(s) during the preparation of the splicing extract may promote its binding to any RNA molecule.

The investigation of the binding specificity of Uss5p, Uss1p and Prp24p shows that all three polypeptides possess a clear preference for the U6 RNA. Very little, if any, interaction was detected between these proteins and the U2 or actin RNAs.

# CHAPTER SEVEN

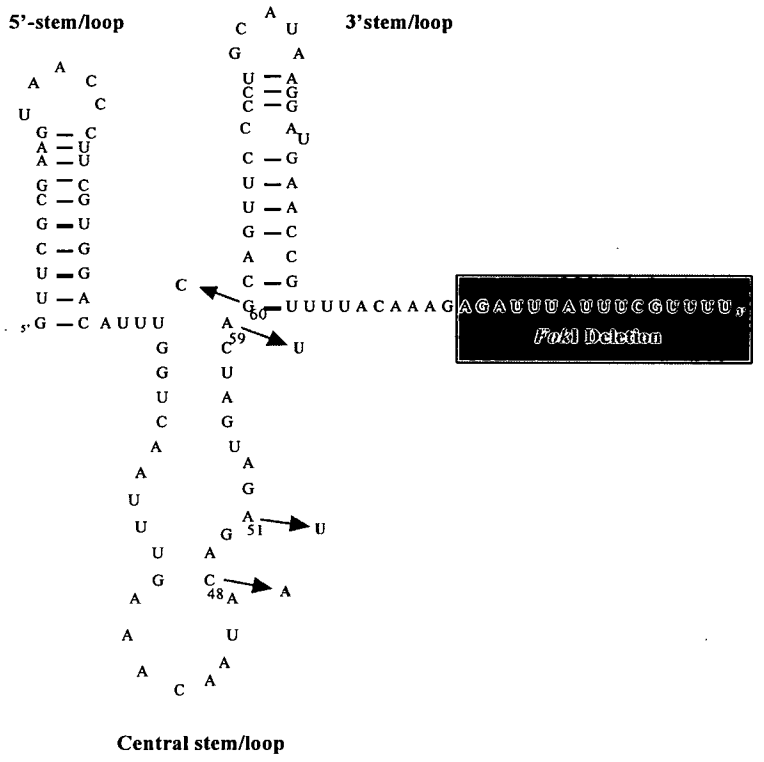
## Effect of U6 snRNA Mutations on Protein Interactions

## INTRODUCTION

SnRNPs and spliceosomes assembly can be affected by U6 snRNA point mutations. In particular, the U6 snRNA C48A and G60C mutations block the *in vitro* splicing reaction prior to step I, while the A51U and A59U mutations block *in vitro* splicing at step II (Fabrizio *et al.*, 1989, 1990 and this work). *In vivo*, all these mutations are associated with a lethal phenotype (Madhani and Guthrie, 1992; McPheeters, 1996). The effect of these mutations on protein contacts was analysed *in vitro* by UV-crosslinking. Hopefully this will bring greater insight on the implication of these U6 snRNA nucleotides on protein interactions and/or it will give a preliminary map of the binding sites (s) of the polypeptides onto the U6 snRNA.

U6 snRNPs reconstituted with a truncated U6 RNA (U6 RNA/*Fok*I) lacking the last 18 nucleotides possess a reduced splicing activity (Fabrizio *et al.*, 1989 and this work). The implication of the U6 snRNA 3' tail on protein contacts was investigated by UV-crosslinking of proteins to the shorter U6 RNA/*Fok*I.

The U6 snRNA mutations and truncation studied in this work are shown below on the free U6 snRNA structure proposed by Fortner *et al.* (1994):




## 7.1 U6 SnRNA Site Directed Mutagenesis and Testing U6 SnRNA Mutants Functionality

For the production of C48A, A59U and G60C mutant U6 RNAs, the wild-type U6 snRNA-encoding gene from the phagemid pUCU6f1 (Table 9) was used as mutagenic template DNA (section 2.4.3.12). The mutagenic oligodeoxynucleotides T1849, T1851 and T1850 (Table 8) were employed for the synthesis of respectively C48A, A59U and G60C mutated U6 snRNA genes. Plasmids harbouring the mutated U6 snRNA genes were screened by sequencing (section 2.4.3.13). The plasmid DNA (pU6A51U) containing the gene encoding the A51U mutant U6 snRNA was a gift from I. Kelly (University of Cambridge). The U6 snRNA/*Fok*I truncation was produced by run-off transcription of pUC12U6 plasmid template linearised with *Fok*I.

The functionality of U6 snRNPs reconstituted with the mutated or truncated U6 RNAs was assayed by *in vitro* splicing reaction (Figure 32). When U6 snRNPs were reconstituted with the wild-type U6 RNA, nearly all the initial splicing activity of the extract was recovered on the basis of lariat-intron-exon 2 and lariat-intron signals, while only 30% of the initial spliced RNA was detected (compare lanes 1 and 3). The mRNA was probably unstable. As expected, no splicing activity was detected with U6 RNA mutants C48A and G60C (lanes 4 and 7). The U6 RNA A51U mutation completely blocked the second trans-esterification step (Step II) of the splicing reaction (lane 5), whereas inhibition was less with the A59U mutation (a small fraction of lariat-intron and spliced RNA was detected) (lane 6). Only 50% of the splicing activity detected with the wild-type U6 RNA was recovered with the truncated U6 RNA (compare lanes 3 and 8).

## 7.2 Proteins Contacting the Mutated and Truncated U6 SnRNAs

U6 snRNPs were reconstituted with the wild-type or the mutated/truncated U6 RNA randomly radiolabelled with $^{32}$P-αATP. After UV-crosslinking, proteins were immunoprecipitated with crude anti-Prp24 antibodies, then resolved on a 15 % SDS-PAGE gel (Figure 33). For an unknown reason, a diffuse signal was observed instead of a 16 kD protein UV-crosslinked to the wild-type U6 RNA (compare Figure 25, lane 3 and Figure 33, lane 3). The C48A U6 RNA mutation and the *Fok*I truncation dramatically reduced the binding of Prp24p (with a more substantial reduction with U6 RNA/*Fok*I truncation), while Uss1p was not detected anymore (Figure 33, lanes 4 and 8). The lack of the Uss1p binding might either reflect loss of protein-U6 RNA contact or be a consequence of loss of Prp24p interaction. Other U6 RNA mutations tested did not affect protein-U6 RNA contacts (Figure 33, compare lane 3 with lanes 5-7). No polypeptides were detected on a 6% gel (data not shown).

The profile of proteins UV-crosslinked to the wild-type, mutated or truncated U6 RNA ($^{32}$P-αATP) and immunoprecipitated with the crude anti-Uss1 and anti-Prp8 antibodies

**Figure 32. Splicing activity of U6 snRNPs reconstituted with wild-type, mutated or truncated U6 RNA.**

U6 snRNPs were reconstituted with wild-type, mutated or truncated U6 snRNA (lanes 3-8) in a 100 µl reaction. An aliquot (9 µl) was taken off and 1 µl substrate RNA (20 000 cpm) was added to assay for the splicing activity. Splicing reaction only was carried out (with preincubation step at 30°C for 30 min) in lane 1; endogenous U6 snRNA was degraded by RNase H in the presence of 140 nM of U6 knockout oligo and no U6 RNA was added (lane 2). After splicing, RNAs were recovered and resolved on a 7% polyacrylamide/urea gel. IVS-E2: lariat-intron exon 2; IVS: lariat excised intron; pre-mRNA: actin substrate RNA p283 (Table 9); E1-E2: spliced RNA; U6 RNA: U6 snRNA transcribed with the T7 RNA polymerase and randomly radiolabelled ($^{32}$P-αATP); E1: exon 1.

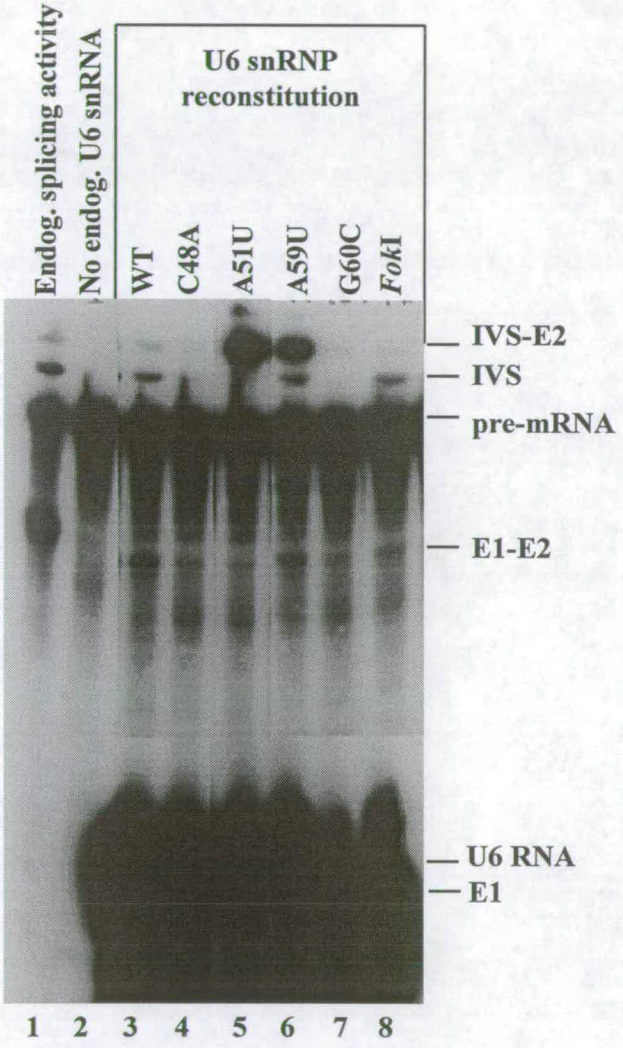
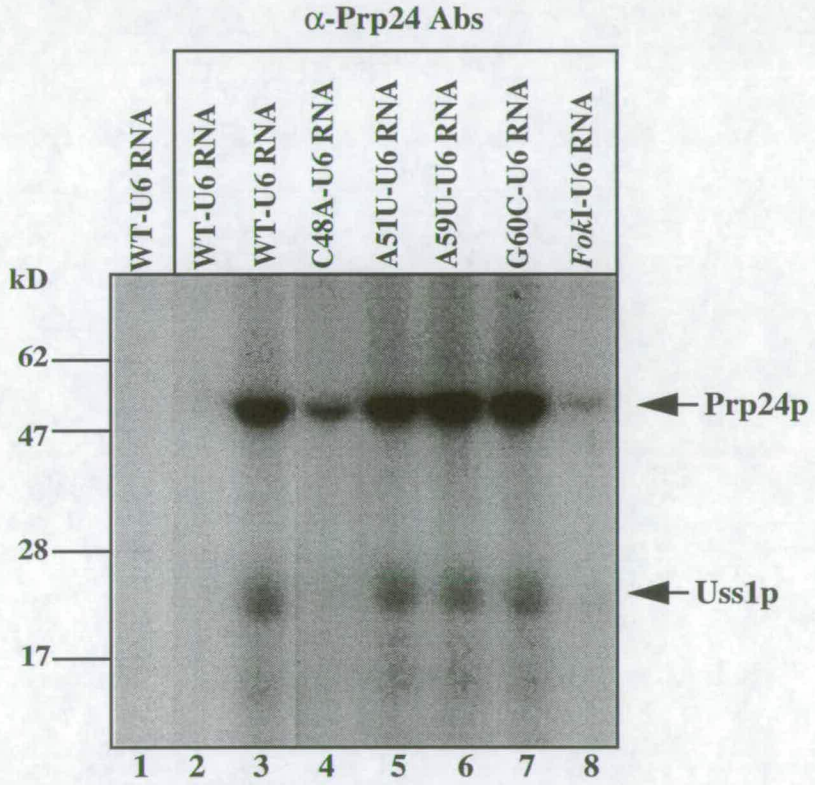

Endog. splicing activity
No endog. U6 snRNA
U6 snRNP reconstitution
WT
C48A
A51U
A59U
G60C
FokI
IVS-E2
IVS
pre-mRNA
E1-E2
U6 RNA
E1
1 2 3 4 5 6 7 8

**Figure 33. Analysis of the effect of U6 snRNA mutation and truncation on protein-U6 RNA interactions.**

U6 snRNPs were reconstituted with 5 nM of randomly radiolabelled wild-type, mutated or truncated U6 RNA ($^{32}$P-αATP) in a 100 µl reaction. Samples were UV-irradiated and protein complexes co-immunoprecipitated with the crude anti-Prp24 antibodies (batch 1016). Reaction sample was not UV-irradiated in lane 2 and preimmune serum was incubated with proteins in lane 1. Immunoprecipitated protein complexes were washed with NTN buffer containing 150 mM NaCl, after RNase treatment half of the precipitated proteins were loaded on a 15% SDS-PAGE and the other half was loaded on a 6% SDS-PAGE (this gel was not shown as no signal was detected).


α-Prp24 Abs
WT-U6 RNA
WT-U6 RNA
WT-U6 RNA
C48A-U6 RNA
A51U-U6 RNA
A59U-U6 RNA
G60C-U6 RNA
FokI-U6 RNA
kD
62
47
28
17
Prp24p
Uss1p
1 2 3 4 5 6 7 8

is shown in Figure 34. The comparison of proteins UV-crosslinked to the wild-type and to the mutated/truncated U6 RNAs suggested that, with the exception of the 90 kD polypeptide (6% gel, lane 6), proteins immunoprecipitated with anti-Uss1 antibodies were less efficiently UV-crosslinked to C48A mutant U6 RNA. In particular the 45 kD polypeptide (which probably reflects the binding of Prp24p to the U6 RNA; see chapter 5) was hardly detectable (compare lanes 4 and 6). With the exception of the 45 kD and 90 kD bands, the diminishion of the intensity of the protein signals was proportional to that of Uss1p. Therefore, the reduction of the UV-crosslinking efficiency might just be a consequence of a weaker Uss1p interaction. The C48A U6 RNA mutation weakly affected proteins UV-crosslinked and immunoprecipitated with anti-Prp8 antibodies; only signals detected on the 15% gel were reduced (compare lanes 5 and 7).

The A51U U6 RNA mutation had no obvious effect on proteins UV-crosslinked and immunoprecipitated with anti-Uss1 or with anti-Prp8 antibodies (Figure 34, compare lane 4 with 8 and lane 5 with 9).

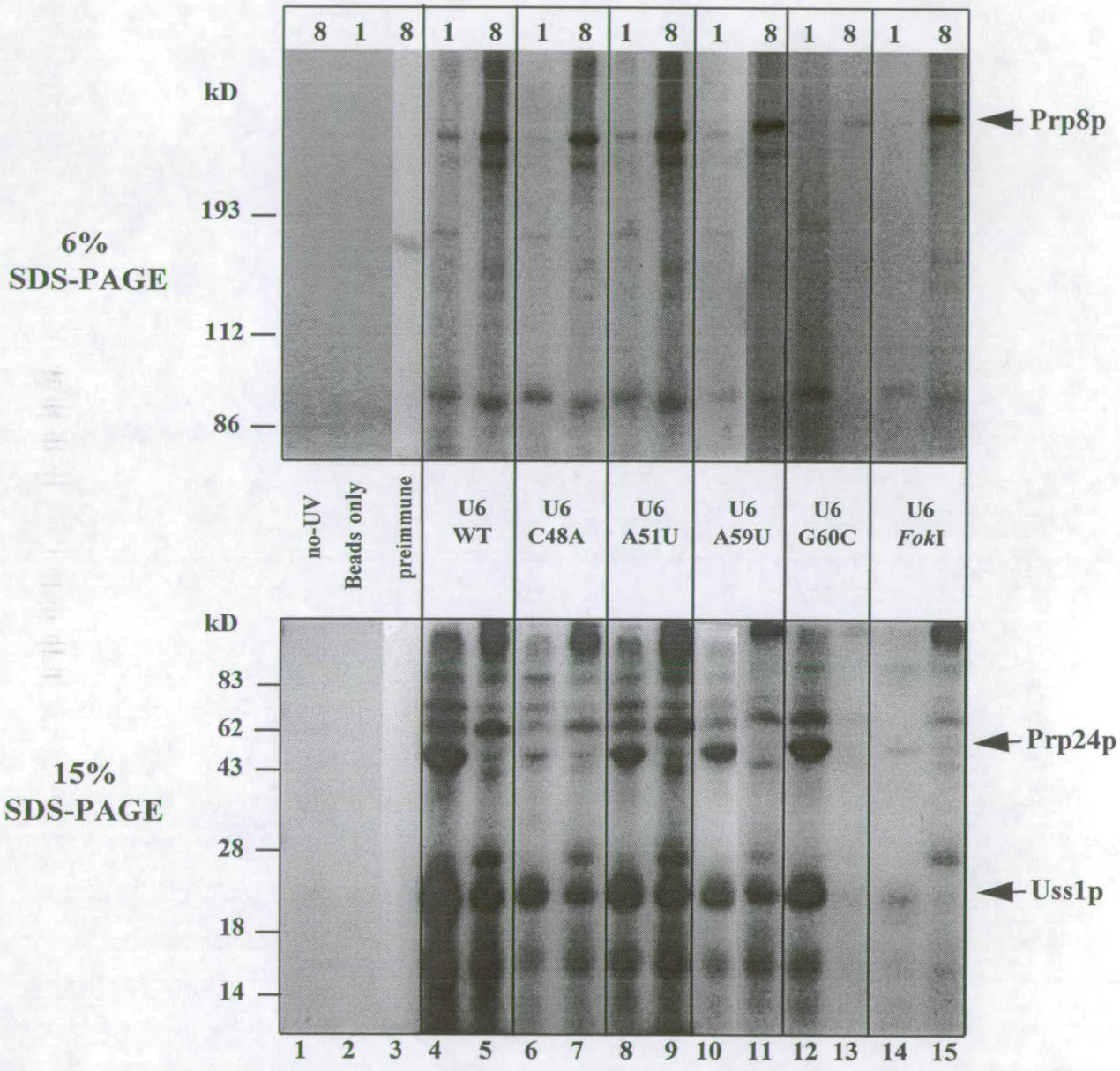
The A59U U6 RNA mutation partially affected the binding of proteins contacting the U6 RNA and immunoprecipitated with anti-Uss1 or with anti-Prp8 antibodies, in that protein signals were reduced (Figure 34, compare lane 4 with 10 and lane 5 with 11).

The G60C U6 RNA mutation prevented the direct interaction of the 70-80 kD protein (Figure 34, 15% gel, compare lane 4 with 12) and the binding of Prp8p detected after immunoprecipitation with the crude anti-Uss1 antiserum (6% gel, compare lane 4 with 12). No proteins co-immunoprecipitated with anti-Prp8 antibodies were detected on the 15% gel, only a weak Prp8p band was observed on the 6% gel (lane 13). The lack of proteins UV-crosslinked to the G60C mutant U6 RNA and co-immunoprecipitated with anti-Prp8 antibodies might be a consequence of the loss of Prp8p interaction.

For the U6 *Fok*I truncation, with the exception of the ~90 kD signal, hardly any proteins were UV-crosslinked and co-immunoprecipitated with anti-Uss1 antibodies (Figure 34, lane 14). No effect was noticed for the proteins immunoprecipitated with

**Figure 34. Proteins UV-crosslinked to the mutated and truncated U6 RNA.**

U6 snRNPs were reconstituted with 5 nM of randomly radiolabelled wild-type, mutated or truncated U6 RNA ($^{32}$P-αATP) in a 100 µl reaction. A 9 µl aliquot was assayed for splicing activity (Figure 32), and with the remaining 91 µl, proteins were UV-crosslinked and immunoprecipitated with either the crude anti-Uss1 antiserum or with the crude anti-Prp8 antiserum. As controls, a reaction sample was not UV-irradiated (lane 1); samples were incubated with PAS beads (lane 2) or with preimmune serum (lane 3). Immunoprecipitated protein complexes were washed with 150 mM NaCl, after RNase treatment half of the precipitated proteins was loaded on a 15% SDS-PAGE gel, the remaining half was loaded on a 6% SDS-PAGE gel.
1: proteins immunoprecipitated with anti-Uss1 antibodies; 8: proteins immunoprecipitated with anti-Prp8 antibodies.


8 1 8 1 8 1 8 1 8 1 8 1 8 1 8
kD
193
112
86
6%
SDS-PAGE
Prp8p
no-UV
Beads only
preimmune
U6 WT
U6 C48A
U6 A51U
U6 A59U
U6 G60C
U6 FokI
kD
83
62
43
28
18
14
15%
SDS-PAGE
Prp24p
Uss1p
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

anti-Prp8 antibodies and resolved on the 6% gel, whereas only two weak signals (28 kD and ~60 kD) remain on the 15% gel (lane 15).

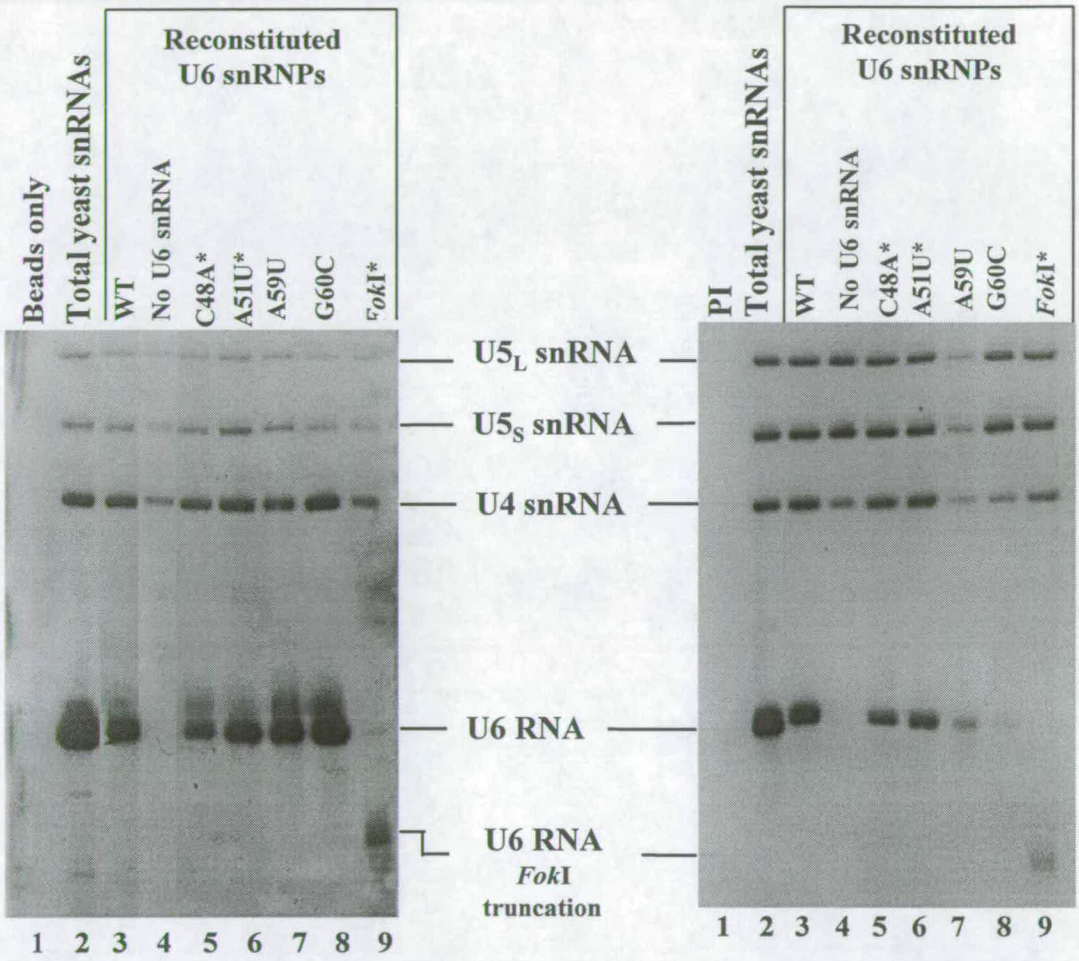
On the basis of these UV-crosslinking experiments, it was unclear whether the absence of protein signal reflected the loss of protein-U6 RNA interaction and/or a detrimental effect on snRNP formation/stability. The effect of U6 RNA mutations and truncation on snRNP formation was therefore further investigated as follows. After U6 snRNP reconstitution, the crude anti-Uss1 or anti-Prp8 antisera were used to immunoprecipitate the snRNP particles. U6 snRNPs, U4/U6 snRNPs and to a lesser extend U4/U6.U5 tri-snRNPs were co-immunoprecipitated with the crude anti-Uss1 antiserum; essentially U5 snRNPs and tri-snRNPs were co-immunoprecipitated with the crude anti-Prp8 antibodies (no substrate RNA was added in the reconstitution reaction). The co-immunoprecipitated U4, U5 and U6 snRNAs were recovered and detected by snRNA blotting (section 2.4.4.3) (Figure 35, Panels A and B).

The precipitation of snRNAs was antibody specific (Figure 35, Panels A and B, lane 1) and the depletion of the endogenous U6 snRNA by RNase H degradation was total (lane 4). By comparison with the sample containing wild-type U6 RNA, in the absence of endogenous U6 snRNA, lower amounts of U4 snRNA and similar amounts of U5 snRNA were co-immunoprecipitated with anti-Uss1 and anti-Prp8 antibodies (compare lanes 3 and 4). Uss1p and Prp8p interactions with U4 snRNA and U5 snRNA suggest that these snRNAs remain associated with each other upon degradation of U6 snRNA.

SnRNAs detected after U6 snRNP reconstitution and immunoprecipitation with anti-Uss1 antibodies are shown in Figure 35 (Panel A). The comparison of U4 snRNA and U6 RNA immunoprecipitated after U6 snRNP reconstitution with wild-type or mutated/truncated U6 RNA indicated that slightly less C48A U6 RNA but as much U4 snRNA were co-immunoprecipitated (compare lanes 3 and 5). A similar variation in the amount of wild-type and C48A U6 RNA was detected in the immunoprecipitation supernate (data not shown), suggesting that this difference was probably due to the fact that only one probe annealed to C48A U6 RNA while two hybridized to the wild-type

**Figure 35. Study by snRNA blotting of snRNP formation.**

A reconstitution reaction (50 µl) was carried out with 10 nM of unlabelled wild-type, mutated or truncated U6 RNA (transcribed from pUC12U6 (Table 9) linearized with *Dra*I or *Fok*I). The endogenous U6 snRNA was degraded by RNase H degradation in sample 4. SnRNAs were co-immunoprecipitated with the crude anti-Uss1 (Panel A) or anti-Prp8 (Panel B) antisera. As controls, samples were incubated with preimmune serum (Panel B, lane 1) or with PAS beads only (Panel A, lane 1). The immunoprecipitated protein-RNA complexes were washed with 150 mM NaCl, snRNAs were recovered, resolved on a 6% polyacrylamide/urea gel and transferred by snRNA blotting onto a membrane. U4, U5 and U6 snRNAs were detected by hybridization (section 2.4.4.3) with end-labelled oligodeoxynucleotides (Table 8) 483A, 485A, Taq6A and V5578 respectively. Two oligodeoxynucleotides (Taq6A and V5578) were used to detect U6 snRNA since the former do not anneal to C48A and A51U mutant U6 RNAs and the second do not anneal to the truncated U6 RNA. Mutant U6 RNAs recognized by only one probe are indicated by an asterisk. Total yeast snRNAs were prepared from 5 µl of splicing extract treated with proteinase K, extracted with phenol/chloroform and precipitated with ethanol. The RNA pellet was resuspended in 2 µl of water, an equal volume of 100% formamide loading dye was added, after boiling, the sample was loaded on the same 6% polyacrylamide/urea gel.




SnRNAs co-immunoprecipitated with crude anti-Uss1 Abs

SnRNAs co-immunoprecipitated with crude anti-Prp8 Abs

A

B

U6 RNA. No effect was detected with A51U, A59U and G60C U6 RNA mutations (compare lane 3 with respectively lanes 6, 7, and 8). Less U6 RNA and U4 snRNA were immunoprecipitated with anti-Uss1 antibodies after U6 snRNP reconstitution with the truncated U6 RNA (compare lane 3 with 9). Little difference was observed between the amount of wild-type and truncated U6 RNAs detected in the immunoprecipitation supernate (data not shown). The lack of the last 18 nucleotides was detrimental to the formation of U6 snRNP, U4/U6 snRNP and maybe tri-snRNP. Alternatively the removal of the last 18 nucleotides affects the association of Uss1p with snRNP particles.

The snRNAs co-immunoprecipitated with anti-Prp8 antibodies after U6 snRNP reconstitution are shown in Figure 35 (Panel B). Comparison of U4 snRNA and U6 RNA immunoprecipitated after U6 snRNP reconstitution with wild-type U6 RNA or with mutated/truncated U6 RNA suggested that less C48A mutant U6 RNA was immunoprecipitated (compare lane 3 with 5). A similar variation was detected in the immunoprecipitation supernate, indicating that this difference was presumably due to the difference in the homology of the probes used to detect the U6 RNAs. The C48A U6 RNA mutation did not affect tri-snRNP formation. The A51U U6 RNA mutation did not have a detrimental effect on tri-snRNP formation (compare lane 3 with 6). Less U4, U5 and U6 snRNAs were recovered after immunoprecipitation of snRNP containing the A59U mutant U6 RNA (lane 7), an unfortunate RNase degradation was responsible for this nonspecific decrease of snRNA signals. This experiment has been reproduced several times and snRNP particles formation was not affected by A59U U6 RNA mutation. A smaller amount of U4 snRNA and G60C mutant U6 RNA was immunoprecipitated with anti-Prp8 antibodies after U6 snRNP reconstitution (compare lane 3 with 8). However, a similar quantity of the G60C and wild-type U6 RNA was added to the reaction (data not shown) suggesting that this mutation was detrimental to tri-snRNP formation and/or Prp8p association with the tri-snRNP. A reduced amount of U4 snRNA and truncated U6 RNA was immunoprecipitated (compare lane 3 with 9). The snRNA analysis of the immunoprecipitate supernate (data not shown) indicated that the low recovery of truncated U6 RNA was not only due to the difference of

**Table 15. Effects of U6 snRNA mutations and truncation on U6 snRNA-protein interactions.**

After immunoprecipitation the recovery of mutated/truncated U6 RNAs was compared to that of wild-type and indicated as follows: +/- for very low; ++ for weaker; +++ for as good as wild-type U6 RNA.

| | **IP with crude α-Prp24 Abs (U6 snRNP)** | **IP with crude α-Uss1 Abs (U6 snRNP, U4/U6 snRNP, U4/U6.U5 snRNP)** | | **IP with crude α-Prp8 Abs (U5 snRNP, U4/U6.U5 snRNP)** | |
|---|---|---|---|---|---|
| ***U6 snRNA mutants*** | ***Effect on UV-crosslinking*** | ***Effect on UV-crosslinking*** | ***U6 snRNA Co-IP*** | ***Effect on UV-crosslinking*** | ***U6 snRNA Co-IP*** |
| ***C48A*** | Much reduced Prp24p signal<br>Loss of Uss1p contact | All signals reduced with Prp24p nearly undetectable | ++ | No significant effect | +++ |
| ***A51U*** | No significant effect | No significant effect | +++ | No significant effect | +++ |
| ***A59U*** | No significant effect | No significant effect | +++ | No significant effect<br>Loss of signal at ~43 kD | +++ |
| ***G60C*** | No significant effect | Loss of a band at 70-80 kD and Prp8p signals | +++ | No proteins on a 15% SDS-PAGE gel<br>Only a weak Prp8p signal remains on 6% SDS-PAGE gel | +/- |
| ***FokI truncation*** | Much reduced Prp24p signal<br>Loss of Uss1p contact | All signals dramatically reduced, only a ~90 kD band remain | +/- | Signals detected on 15% gel (only) dramatically reduced | +/- |

homology of the probes used to detect the shorter and the wild-type U6 RNAs. Therefore, the removal of the last 18 nucleotides of U6 RNA affects tri-snRNP formation/stability and/or Prp8p association with this particle.

The effects of U6 snRNA mutations and truncation are summarised in Table 15.

## 7.3 Discussion

This work showed that U6 and/or U4/U6 snRNP and U4/U6.U5 tri-snRNP formation, assayed by immunoprecipitation of snRNPs and snRNA blotting, was not affected by A51U and A59U U6 snRNA mutations. These mutations did not interfere with protein-U6 RNA contacts detected before spliceosome formation (no substrate RNA was added) as demonstrated by UV-crosslinking and immunoprecipitation. Substitution of A to U, C or G, at U6 snRNA position 51 or 59, blocks the second step of the splicing reaction (Fabrizio *et al.*, 1991). Therefore, as the results presented in this chapter suggest, snRNPs and spliceosome assembly occur as in the wild-type situation. These data also indicate that, before spliceosome formation, U6 snRNA nucleotides A51 and A59 nucleotides are not essential for the direct binding of proteins co-immunoprecipitated with anti-Uss1, anti-Prp24 or anti-Prp8 antibodies.

More dramatic effects were observed with the C48A U6 RNA mutation. It affects:
- U6, U4/U6 snRNP and U4/U6.U5 tri-snRNP formation (very mild effect).
- The binding of proteins (Prp24p and Uss1p) to the U6 snRNA in the U6 snRNP and immunoprecipitated with the anti-Prp24 antiserum.
- The UV-crosslinking of proteins (Prp24p and Uss1p) to U6 snRNA in the U6 snRNP and/or in the U4/U6 snRNP and immunoprecipitated with the anti-Uss1 antiserum.

Despite the reduced binding of Prp24p to C48A mutant U6 RNA, Prp24 fulfilled its function since U4/U6 snRNPs and tri-snRNPs form (Fabrizio *et al.*, 1990; and this work), thus two hypotheses are suggested:

- Although the binding of Prp24p to U6 snRNA was dramatically reduced, the remaining contact was sufficient to drive U4/U6 snRNPs formation.
- Prp24p-U6 RNA UV-crosslinking did not reflect a functional interaction therefore the reduction of the radioactive signal did not correlate with an alteration of Prp24p's function.

The binding of some proteins (in particular Prp24p) to U6 RNA was affected by the C48A mutation (this study). Despite these detrimental effects, U4 snRNA, C48A U6 RNA and U5 snRNA seem to interact with each other to form U4/U6 snRNP and U4/U6.U5 tri-snRNP which were immunoprecipitated with anti-Uss1 and anti-Prp8 antisera. These particles were similar to those containing the wild-type U6 RNA (in respect to snRNP containing the wild-type U6 RNA and immunoprecipitated with anti-Uss1 and anti-Prp8 antisera) (this study). Fabrizio *et al.* (1990) reported that *in vitro,* no active spliceosome form with the C48A U6 RNA mutant. Presumably, the tri-snRNPs containing the mutated U6 RNA were unable to associate with the prespliceosome, or further progression toward active spliceosome was blocked. The C48A U6 RNA mutation is located in the highly conserved sequence $A_{47}CAGA_{51}$ which interacts with the pre-mRNA 5' splice site (Lesser *et al.*, 1993; Kandels-Lewis *et al.*, 1993). This sequence has been proposed to start base-pairing with the 5' splice site before disruption of U6 snRNA base-pairing with the U4 snRNA (Konforti *et al.*, 1994). The interaction between this U6 snRNA sequence and the 5' splice site may also be required to disrupt U4/U6 snRNAs interaction (Li *et al.*, 1996). Thus, the C48A mutation may destabilize the U6 snRNA-5' splice site base-pairing and prevent U4/U6 snRNP dissociation, blocking spliceosome maturation. Kandels-Lewis *et al.* (1993) reported that the splicing defect associated with C48A mutation was suppressed by a complementary mutation in the intron at position 5 (G5C). This complementation never restored splicing activity to a wild-type level, suggesting that the U6 snRNA C48A mutation affects the interaction of other factor(s) with the U6 snRNA. Prp24p is suspected to disturb the U4/U6 snRNA stem II and facilitate the action of other splicing factor(s) required for U4/U6 snRNP dissociation prior to the first step of splicing (Ghetti *et al.*, 1995). Vaidya *et al.* (1996) have reported a genetic interaction between

Prp24p and Prp21p (a component of the U2 snRNP) suggesting that Prp24p is associated with the immature spliceosome. The results presented in this work suggest that the C48A U6 RNA mutation destabilizes Prp24p binding to the U6 snRNA in the U6 snRNP. Taking these data together one can speculate that C48A U6 snRNA mutation destabilizes (U6 snRNA-5'splice site) and (U6 snRNA-Prp24p) interactions. Consequently, U4 snRNA remains in the spliceosome and blocks spliceosome activation.

U6 and U4/U6 snRNPs formed with G60C mutant U6 RNA (with respect to snRNPs immunoprecipitated with anti-Uss1 antiserum), also protein contacts were only slightly affected (this study). This mutation prevented the detection of proteins UV-crosslinked to the G60C U6 RNA and immunoprecipitated with anti-Prp8 antiserum, suggesting that the G60C substitution had a detrimental effect on tri-snRNP formation and/or Prp8p association with this particle.

The U6 snRNA 3' terminal domain (nucleotides 81-112) is poorly conserved in sequence but is strictly conserved in size (between 29 nt and 32 nt in all species). Unexpectedly, alterations in the length of this domain were associated *in vivo* with mild phenotypes (Madhani *et al.*, 1990) and *in vitro* did not dramatically affect the splicing reaction (Fabrizio *et al.* 1989 and this work). Further 3' end truncations of the U6 snRNA (U6/*Bcl*I: lacks the last 22 nucleotides, U6/*Eco*NI: lacks the last 32 nucleotides) do not restore splicing activity of the extract after U6 snRNPs reconstitution (Fabrizio *et al.*, 1989). Probably, the removal of the last 18 nucleotides is the maximum truncation allowed to maintain U6 snRNA's function. On a nondenaturing gel, U4/U6 snRNPs formed with the *Fok*I truncated U6 RNA possess a similar mobility to particles containing the wild-type U6 RNA (Fabrizio *et al.*, 1989). Also, presumably, tri-snRNPs and spliceosomes have formed since splicing occurs. UV-crosslinking and immunoprecipitation experiments have shown that this truncation affected proteins contacting the U6 snRNA in the U6 snRNP and in the U4/U6 snRNP particles (this study). In particular Prp24p and Uss1p binding was dramatically reduced. In the tri-snRNP Prp8p signal remained unchanged whereas Uss1p contact was not detected

anymore. These results suggest that either Uss1p and Prp24p were UV-crosslinked to the 3' tail of the U6 RNA or that the 3' end was required for Uss1p and Prp24p to contact the U6 RNA elsewhere. Alternatively, other protein(s) may interact with the last 18 nucleotides of U6 RNA and help Uss1p and Prp24p to bind U6 snRNA elsewhere. Fortner *et al.* (1990) observed that in their yeast extract the U6 snRNA was degraded by a 3' nuclease activity producing a 90 nt and a 60 nt U6 snRNA forms (the nuclease was probably paused by the U6 snRNA 3' stem loop (nt 61-87), suggesting that the 3' tail was not strongly protected from nuclease degradation by protein(s). This region is more likely to be free or weakly associated with proteins. According to the U6 snRNA secondary structure proposed by Fortner *et al.* (1990), the last 18 nucleotides are not involved in any RNA-RNA interactions in U6 snRNP or U4/U6 snRNP. Therefore, presumably by being free, the U6 snRNA 3' tail may act as a flexible arm and guide and/or stabilize protein(s) onto the U6 snRNA.

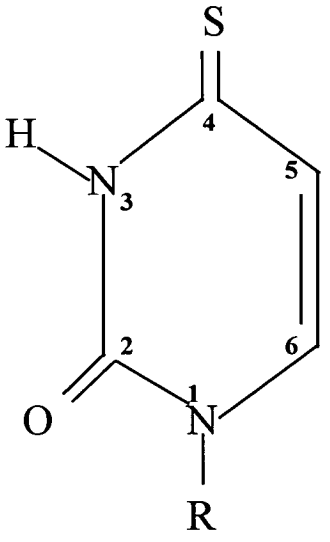
# CHAPTER EIGHT

## Analysis by Site Specific UV-Crosslinking of Proteins Contacting the U6 snRNA

## INTRODUCTION

In order to map the binding site of proteins to the U6 RNA, a single photoactivatable reagent (4-thiouridine) was incorporated into a specific position of the U6 snRNA. Protein-U6 snRNA contacts were subsequently investigated by UV-crosslinking.

4-Thiouridine (4-thioU) is a uridine analog in which the keto oxygen at position 4 of the pyrimidine ring is replaced by a sulphur atom (Figure 36).




R = ribosyl

**Figure 36. Structure of the 4-thiouridine**

Due to the replacement of the O by a S, the C-S bond is twice the length of the C-O bond (0.8 Å instead of 0.4 Å). In the case of a Watson-Crick [A-$S^4$U] base-pair the NH...O bond (2.9 Å) is replaced by an NH...S bridge of 3.4 Å, corresponding to a 15-20% increase in bond length. Despite these structural variations, it has been shown that low substitution levels appear to have minor effects on polynucleotide conformation and stability (Favre, 1990). The incorporation of a single 4-thioU in the U6 snRNA molecule did not interfere with its functionality (this chapter). The 4-thioU possesses

base-pairing properties similar to those of unmodified uridine and is very efficiently photoactivatable by wavelengths above 320 nm which keep undesirable side reactions at a minimum. RNAs containing 4-thiouridine have been widely used to study RNA-RNA and RNA-protein interactions since this photoactivatable base possesses a high photoreactivity with both amino acids and nucleotides (Sontheimer, 1994).

For two main reasons the binding of proteins to the U6 snRNA nucleotide $U_{54}$ has been investigated:

- $U_{54}$ is situated after the invariant U6 snRNA sequence $A_{47}CAGAGA_{53}$ located in a region suspected to be part of the catalytic centre of the spliceosome (Madhani *et al.*, 1992, 1994; Lesser *et al.*, 1993)
- Following the method described in section 2.4.4.5, the site of incorporation of the photoactivatable chemical must be at an UpG position. Since a G residue follows $U_{54}$, the incorporation of a 4-thioU will not necessitate mutation of the U6 snRNA sequence.

## 8.1 Production of a U6 RNA Containing a Radiolabelled 4-Thiouridine at Position $U_{54}$

The method described in section 2.4.4.5 was followed and both transcription DNA templates used to produce the 5' and the 3' halves of the U6 snRNA were amplified by Polymerase Chain Reaction (section 2.4.3.8).
The PCR programme applied was:

| | | |
|---|---|---|
| Denaturation step: | 94°C | 1 min |
| Annealing step: | 50°C | 1 min |
| Amplification step: | 70°C | 2 min |
| N$^0$ of cycles: | 35 | |

For the amplification of the 5' half transcription template, oligodeoxynucleotides S5760 and W3464 were used as 5' and 3' primers respectively and pUC12U6 was utilized as the DNA template. For the amplification of the 3' half transcription template, oligodeoxynucleotides W3463 and S5761 were used as 5' and 3' primers respectively and pU6 was utilized as the DNA template. Two U6 snRNA 3' halves were produced by transcription *in vitro*, one contains at its 5' extremity a 4-thiouridine whereas the other (as a control) possesses a uridine. These 3' U6 snRNA moieties were 5' end radiolabelled and ligated with the 5' U6 RNA half by using oligodeoxynucleotide W3462 as a bridging oligo (Table 8, section 2.4.4.5). The ligation product was resolved on a 8% polyacrylamide-urea gel, detected by autoradiography (section 2.2.2) (data not shown) and gel purifed.

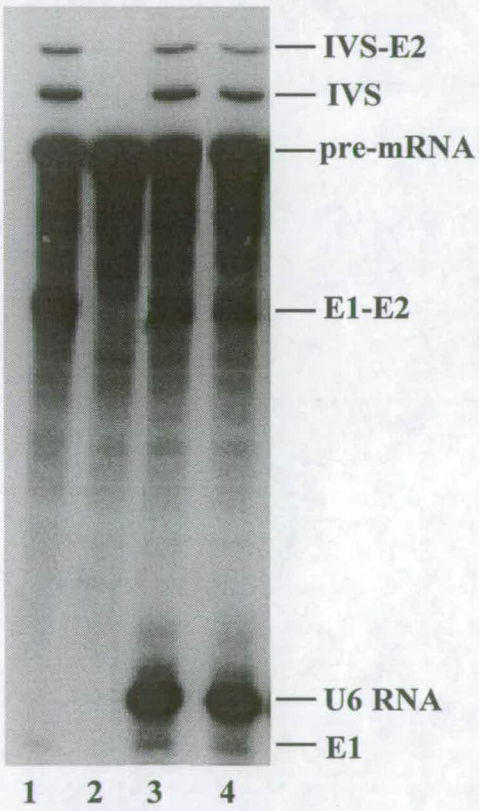
U6 snRNPs were reconstituted with U6 RNA containing a $^{32}$P-U or a $^{32}$P-4-thioU at position 54, and their ability to support splicing was assayed (Figure 37). Eighty to ninety percent of the initial splicing activity was recovered (on the basis of lariat-intron-exon 2 and lariat-intron signals (compare lane 1 with lanes 3 and 4)), indicating that both U6 RNA molecules are functional and that the quantity (unknown) of U6 RNA added allows U6 snRNPs reconstitution with a high efficiency. A slightly lower splicing activity was detected with U6 RNA containing the 4-thioU, however the comparison of the U6 RNA signals indicated that slightly less 4-thioU-containing U6 RNA was added in the reaction.

Attempts to incorporate a radiolabelled 4-thioU or U at U6 snRNA positions $U_{38, 59}$ and $U_{80}$ failed; three main problems were encounted:

- the transcription efficiency of the 3' U6 RNA halves (initiated by UpG or 4-thioUpG) was very low and often not enough RNA was produced, even when different transcription conditions were tested. For an unknown reason there was a high level of abortive transcription.
- the 5' end radiolabelling of the 3' RNA halves was inefficient, which may reflect an inefficient initiation of the transcription with UpG or with 4-thioUpG.

**Figure 37. U6 snRNP reconstitution with U6 snRNA containing a $^{32}$P-U or a $^{32}$P-4thioU at position $U_{54}$.**

A splicing extract (prepared from the yeast strain SC261) was preincubated for 30 min at 30°C before testing its splicing activity (lane 1). No splicing activity was left when endogenous U6 snRNA was degraded by RNase H in the presence of 140 nM U6 knockout DNA oligo (lane 2). U6 snRNP particles were reconstituted with a U6 RNA containing a $^{32}$P-uridine (lane 3) or a photoactivatable $^{32}$P-4-thiouridine (lane 4) at position $U_{54}$. From a 100 µl reconstitution reaction (prepared to UV-crosslink proteins to U6 snRNA; Figure 38), an aliquot (9 µl) was taken to which 1 µl of radiolabelled substrate RNA (20 000 cpm) was added. After splicing, RNAs were recovered and resolved on a 7% polyacrylamide/urea gel. IVS-E2: lariat-intron-exon 2; IVS: excised lariat-intron; pre-mRNA: actin substrate RNA (p283, Table 9); E1-E2: messenger RNA; E1: exon 1; U6 RNA: U6 snRNA produced by transcription *in vitro* with the $T_7$ RNA polymerase. The amount of U6 RNA added to the reaction was unknown, approximately 200 cps (on the Geiger counter) of U6 RNA was added per 100 µl reconstitution reaction.


IVS-E2
IVS
pre-mRNA
E1-E2
U6 RNA
E1
1
2
3
4

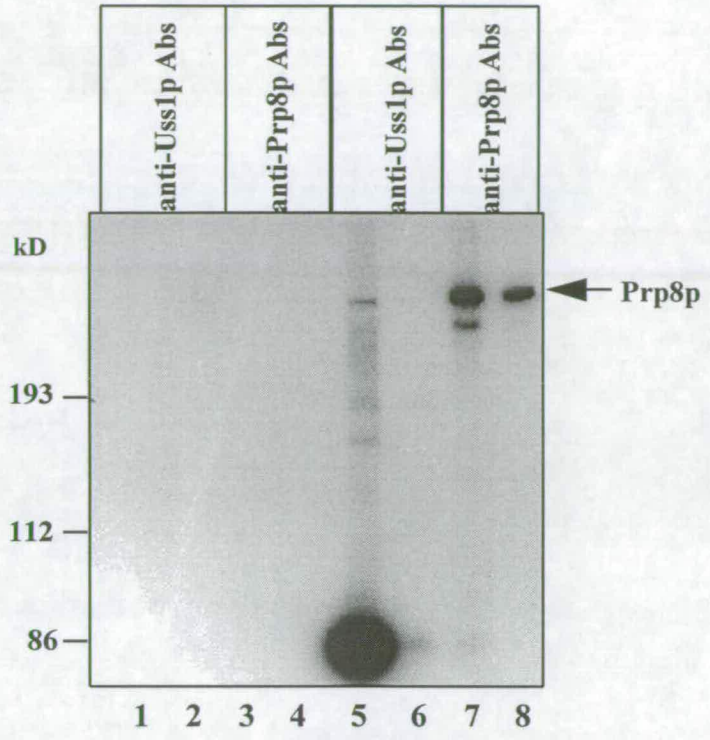
- the ligation of the 5' and the 3' RNA halves was limited, probably because non-templated nucleotides were added by the T7 RNA polymerase at the 3' end of the transcripts. This would consequently reduce the RNA ligation efficiency since a perfect match with the RNA extremities and the bridging oligodeoxynucleotide is required. This 3' end tailing cannot be circumvented, and the best solution would have been to chemically synthesise the 5' half U6 RNA or to generate by RNase H cleavage an RNA with the correct 3' end (Yu and Steitz, 1997; Lapham *et al.*, 1997).

## 8.2 Study of Proteins Interacting with the U6 SnRNA at Position $U_{54}$

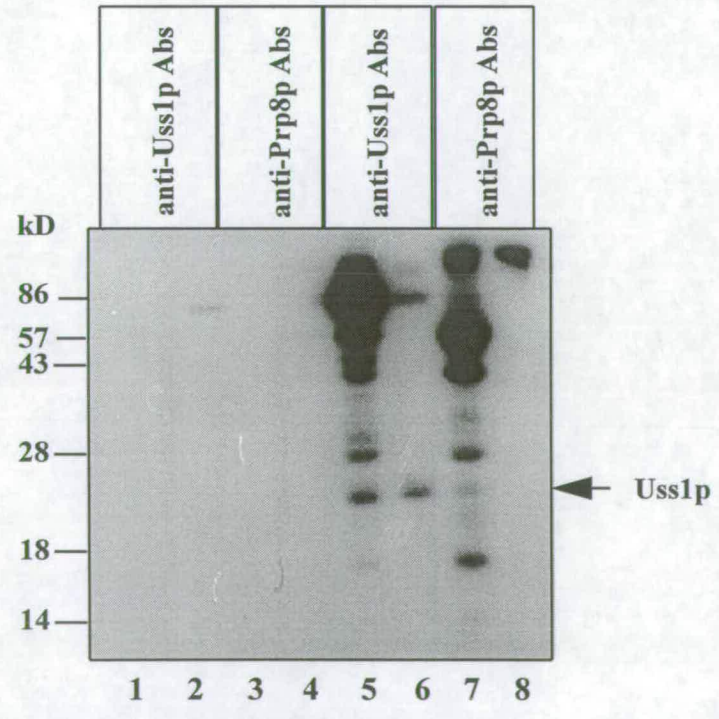
Proteins UV-crosslinked to U6 RNA at position $U_{54}$ and immunoprecipitated with the crude anti-Uss1 or anti-Prp8 antibodies are shown in Figure 38. The UV-crosslinking signals detected were specific for the presence of the photoactivatable 4-thioU (compare lanes 1-4 and lanes 5-8). A protein of ~25 kD was immunoprecipitated with anti-Uss1 antibodies under nondenaturing conditions or when proteins were denatured prior to immunoprecipitation, thus Uss1p contacts U6 snRNA at position $U_{54}$ (lanes 5, 6). At least seven additional polypeptides were immunoprecipitated with the anti-Uss1 antibodies and detected on the 6% and 15% gels. They possessed apparent molecular weights of 28 kD, 40 kD, 60 kD, ~86 kD, 150-170 kD, 190 kD and >200 kD. A protein of >200 kD was immunoprecipitated with anti-Prp8 antibodies under nondenaturing conditions or when proteins were denatured before immunoprecipitation, suggesting that Prp8p contacts U6 RNA at position $U_{54}$ (lanes 7, 8). The >200 kD signal detected after UV-crosslinking and immunoprecipitation with anti-Uss1 antibodies probably reflects the binding of Prp8p to the U6 RNA in the small fraction of tri-snRNPs co-immunoprecipitated with anti-Uss1 antibodies. At least seven additional polypeptides were detected on the 6% and 15% gels after co-immunoprecipitation with anti-Prp8 antibodies. They possessed apparent molecular sizes of 16 kD, 25 kD, 28 kD, 40 kD, ~60 kD and two signals at >220 kD. The 25 kD protein could well be Uss1p since it was shown previously (Figure 26) that it contacts

**Figure 38. Proteins UV-crosslinked to the U6 snRNA at position $U_{54}$.**

U6 snRNP reconstitution reaction (100 μl) was carried out (no substrate RNA was present). The concentration of U6 RNA was unknown, 200 cps (on the Geiger counter) of U6 RNA was added per reconstitution reaction. U6 snRNPs were reconstituted with U6 snRNA containing a $^{32}$P-U (lanes 1-4) or containing a long wavelength photoactivatable $^{32}$P-4-thioU (lanes 5-8) at position $U_{54}$. In lanes 1, 2, 5 and 6, protein complexes were immunoprecipitated with the crude anti-Uss1 antiserum under nondenaturing conditions (lanes 1 and 5) or after denaturation of proteins (lanes 2 and 6). In lanes 3, 4, 7 and 8, protein complexes were immunoprecipitated with anti-Prp8 antiserum under nondenaturing conditions (lanes 3 and 7) or after denaturation of proteins (lanes 4 and 8). Half of the immunoprecipitated proteins was resolved on a 15% SDS-PAGE gel and the other half on a 6% SDS-PAGE gel. Proteins were detected by autoradiography after an overnight exposure.




6% SDS-PAGE




15% SDS-PAGE

**Table 16. Proteins UV-crosslinked to U6 snRNA at position $U_{54}$.**
The intensity of the protein signals was indicated as follows: ++++, very strong; +++, strong; ++, medium; +, weak; +/-, very weak.

the U6 snRNA in the tri-snRNP. The proteins interacting with the U6 RNA at position $U_{54}$ are indicated in Table 16.

| *Protein co-immunoprecipitated with Anti-Uss1 antisera* | | *Proteins co-immunoprecipitated with Anti-Prp8 antisera* | |
|---|---|---|---|
| **15% gel (kD)** | **6% gel (kD)** | **15% gel (kD)** | **6% gel (kD)** |
| $86^{++++}$ | $>200^{+/-}$/Prp8p | | $Prp8p^{+++}$ |
| $\sim60^{++}$ | | $\sim60^{+++}$ | $>200^{++}$ |
| | $190^{+/-}$ | | |
| $40^{++}$ | $150\text{-}170^{+/-}$ | $40^{+++}$ | |
| $28^{++}$ | | $28^{++}$ | |
| $Uss1p^{++}$ | | $25^{+/-}$/Uss1p | |
| | | $16^{++}$ | |

## 8.3 Discussion

Prp8p and Uss1p contact the U6 RNA at position $U_{54}$ in addition to numerous uncharacterized proteins. The site specific UV-crosslinking method was extremely efficient; for a shorter UV exposure time (5 min instead of 25 min) and with only a single photoactivatable nucleotide, many more proteins were UV-crosslinked to $U_{54}$ than were UV-crosslinked to the randomly radiolabelled U6 RNA. Also, the intensity

of the signals was a lot higher since only an overnight exposure of the hypersensitive film was necessary compared to 5-15 days needed to detect proteins contacting the randomly radiolabelled U6 RNA. Unexpectedly numerous protein-U6 RNA interactions were detected at position $U_{54}$ suggesting that snRNP particles were "breathing" leading to dymanic protein-RNA associations.

A substantial variation in the intensity of the protein signals was noticed which may reflect a difference in the UV-crosslinking efficiency between the various polypeptides present in the vicinity of the U6 RNA. In particular, the Uss1p signal was weak, suggesting that although it contacts the U6 RNA at position $U_{54,}$ it may not be the main site of interaction.

Unfortunately technical problems prevented to map the binding site of the proteins at different positions of the U6 RNA. At the time when the experiments were carried out, another technique allowing insertion of a 4-thiouridine at specific positions of the RNA molecule was published (Yu and Steitz, 1997). This new method permitted the introduction of a 4-thiouridine at any site of the RNA independently of its sequence. Also RNA molecules lacking non-templated additional nucleotides were generated by site-specific RNase H cleavage directed by 2'-O-methyl-specific cleavage (Yu and Steitz, 1997). The use of this technique may circumvent the problems encounted in this work. One obvious limitation of this method would be the cost since the synthesis of 2'-O-methyl-containing oligodeoxyribonucleotides is very expensive.

# CHAPTER NINE

## Study of Protein-U6 snRNA Interactions Occuring in the Spliceosome

## INTRODUCTION

The splicing reaction is catalyzed by the spliceosome, which is formed by an ordered interaction of snRNPs and numerous proteins with short conserved sequences of the substrate pre-mRNA (the 5' and 3' splice sites and the branched point). There is considerable evidence that the U6 snRNA is part of the catalytic centre of this large ribonucleoprotein complex (Fabrizio *et al.*, 1990; Madhani *et al.*, 1992, 1994; McPheeters, 1996; Sawa *et al.*, 1991; Sontheimer *et al.*, 1993; Wassarman *et al.*, 1992). Therefore the investigation of U6 RNA-protein interactions occuring in the spliceosome is of great interest and may allow the identification of protein(s) required for splicing. To date only two splicing factors have been reported to interact with U6 snRNA in the spliceosome, Prp16p and Slt22p:

- Prp16p (120 kD), an RNA-dependent ATPase required for the second catalytic step of the splicing reaction (Schwer *et al.*, 1991), has been shown to genetically interact with the U6 snRNA nucleotides 40-46 and also with U2-U6 snRNAs helix I in the spliceosome (Madhani *et al.*, 1994b; McPheeters, 1996).
- Slt22p (246 kD) is an RNA-dependent ATPase and a putative RNA helicase that interacts genetically with a region of U2 snRNA involved in potential U2/U6 helix II. Slt22p possesses an ATPase activity that is stimulated by the presence of U2/U6 snRNAs (Xu *et al.*, 1996), suggesting that U2/U6 hybrid might be a natural binding substrate for this protein.

The method described in section 5.1 was applied to study U6 RNA-protein interactions occuring in the spliceosome. Spliceosome complexes were indirectly isolated by immunoprecipitation with anti-Prp2, anti-poly(A) binding protein (PAB) and anti-Prp8 antibodies:

- Prp2p is a 100 kD protein (Lee *et al.*, 1986) required for promotion of the first step of the splicing reaction (Lin *et al.*, 1987). It is a putative RNA helicase which possesses an RNA-stimulated ATPase activity (Kim *et al.*, 1992). Prp2p interacts transiently with the spliceosome prior to and during step I of splicing (King *et al.*, 1990), consequently, the co-immunoprecipitation of the spliceosome with anti-Prp2 antibodies is inefficient. A

dominant negative mutant Prp2p$^{LAT}$ has been engineered (Plumpton *et al.*, 1994). This mutant protein contains a single amino acid change within the conserved SAT motif that is potentially required for RNA unwinding. Prp2p$^{LAT}$ becomes stalled in the spliceosome complex and blocks the splicing reaction before step I (Plumpton *et al.*, 1994). As this mutated protein is efficiently precipitated with anti-Prp2 antibodies, the spliceosome, poised before step I, can be co-immunoprecipitated with these antibodies.

- The poly(A) binding protein is present in the cytoplasm, with a molecular size of 70 kD, and is also found in the nucleus as two proteolytic forms (53 kD and 13 kD) of the cytoplasmic protein. The 70 kD and 53 kD forms possess the ability to bind poly(A) RNAs with high affinity (Sachs *et al.*, 1986). In the *in vitro* system described in section 5.1, spliceosome formation will be induced by the addition of unlabelled pre-mRNA containing a poly(A) tail, thus PABp will interact with this substrate RNA, and spliceosomes will be co-immunoprecipitated using anti-PAB antibodies (Whittaker and Beggs, 1990).
- Under nondenaturing conditions, anti-Prp8 antibodies co-immunoprecipitate the U5 snRNP, the U4/U6.U5 tri-snRNP and the spliceosome (Brown, 1992) and will be utilized to isolate these complexes.

## 9.1 Proteins Contacting the U6 SnRNA in the Spliceosome

In order to study protein-U6 snRNA contacts occuring in the spliceosome *in vitro*, spliceosome formation was induced by adding unlabelled substrate RNA (RP51A; 1 μM final concentration) following U6 snRNP reconstitution. This substrate possesses a 100 nt-long poly(A) tail and was produced by run-off transcription of the RP51A plasmid linearized with *Bam*HI (Table 9). As a negative control, a mutant RP51A lacking the 5' splice site (RP51A-5'0) and unable to form spliceosomes was added in independent reconstitution reactions. This substrate was produced by run-off transcription of the RP51A-5'0 plasmid linearized with *Bam*HI (Table 9). As an additional control, substrate was ommitted to test for the specificity of binding of PABp.

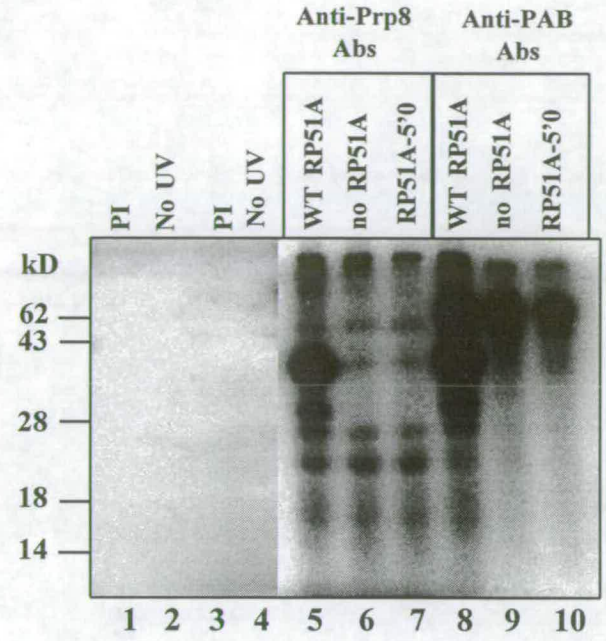
Proteins UV-crosslinked to the U6 RNA ($^{32}$P-αATP) and immunoprecipitated with anti-PAB or anti-Prp8 antisera are shown in Figure 39. Five UV-crosslinked proteins were precipitated with anti-Prp8 antibodies (16 kD, 23 kD, 25 kD, 40 kD and ~60 kD) in the absence of substrate RNA or when no spliceosome form (lanes 6 and 7). When the formation of functional spliceosomes was induced by the addition of unlabelled substrate RNA, the signal at 40 kD was much stronger, and a novel polypeptide (30 kD) was detected (compare lane 5 with lanes 6 and 7). Other protein signals (16 kD, 23 kD, 25 kD and 60 kD) remain unchanged indicating that these proteins contact the U6 RNA at least in the tri-snRNP, however their presence in the spliceosome is uncertain. The intensity of the 40 kD band was greatly increased in a spliceosome-dependent fashion. It may be that the same signal detected, in the absence of substrate or when a nonfunctional RP51A-5'0 was added, reflects a basal level of spliceosomes formed with endogenous pre-mRNA. Alternatively this 40 kD signal may correspond to a different protein interacting with the U6 snRNA in the tri-snRNP, or to a polypeptide which binds weakly to the U6 RNA in the tri-snRNP and more tighly when spliceosomes form. The immunoprecipitation of spliceosomes with anti-PAB antiserum shows that at least five proteins are in close proximity to the U6 RNA (lane 8). They possess apparent molecular weigths of 23 kD (weak signal), ~25 kD, 30 kD, 40 kD and ~70 kD. The 70 kD polypeptide was detected independently of spliceosome formation and remain UV-crosslinked to the U6 RNA and associated with the anti-PAB antiserum after denaturation (data not shown). Thus, this signal reflects the binding of the PAB protein to the U6 RNA.

The mixture of a splicing extract prepared from yeast cells S150-2B overproducing Prp2p$^{LAT}$ with an extract prepared from wild-type yeast cells (SC261) resulted in the accumulation of spliceosomes stalled prior to the first step (i.e. complex I on a native gel; section 2.6.2.2) (Plumpton *et al.*, 1994). Titration experiments showed that the best complex I accumulation was obtained by mixing 2 volumes of Prp2p$^{LAT}$-containing splicing extract with 3 volumes of wild-type yeast extract (data not shown). Splicing extracts were then mixed to such a ratio and U6 snRNPs were reconstituted with randomly radiolabelled U6 RNA ($^{32}$P-αATP). Proteins were UV-crosslinked to U6

**Figure 39. Proteins UV-crosslinked to uniformly radiolabelled U6 RNA ($^{32}$p-αATP) and immunoprecipitated with anti-PAB and anti-Prp8 antibodies.**

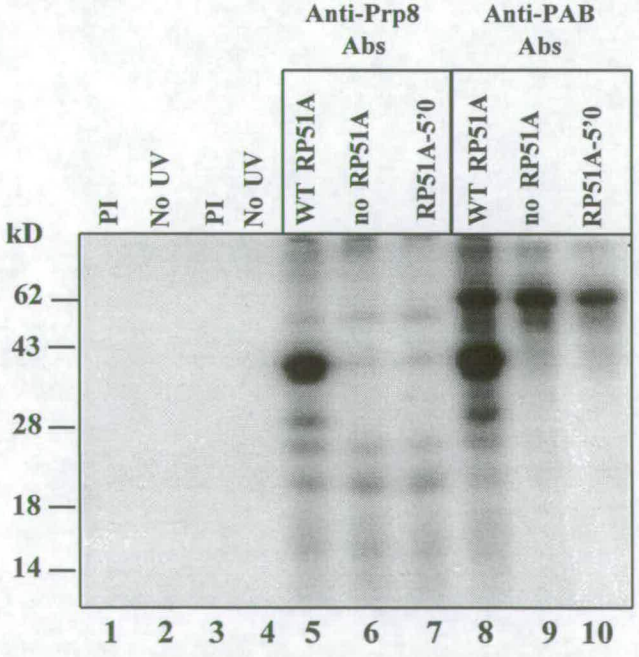
In a splicing extract (50 µl) prepared from wild-type SC261 cells (Table 7), U6 snRNPs were reconstituted (section 2.6.2.3) with 6 nM of randomly radiolabelled U6 RNA ($^{32}$P-αATP). Unlabelled wild-type substrate RP51A (1 µM of final concentration) was then added (lanes 1-5 and lane 8) to induce spliceosome formation. As controls, substrate was omitted (lanes 6 and 9) or an unlabelled mutant RP51A-5'0 pre-mRNA (1 µM) lacking the 5' splice site and unable to form spliceosomes was added (lanes 7 and 10). Samples were UV-irradiated then proteins were incubated with anti-prp8 (lanes 5-7) or anti-PAB (lanes 8-10) antisera. As a control, UV-crosslinking was omitted in samples 2 and 4, or proteins were incubated with preimmune serum (lanes 1 and 3). After washes (150 mM NaCl), proteins were resolved on a 15% SDS-PAGE gel. Two different exposure times of the hypersensitive film are shown (6 h and an overnight exposure).

A




**Overnight exposure of the hypersensitive film**

B




**6h exposure of the hypersensitive film**

RNA in stalled spliceosomes and immunoprecipitated with anti-Prp2 antibodies (Figure 40, Panel A). At least three polypeptides (30 kD, 40 kD and 45 kD) were specifically UV-crosslinked to U6 RNA in the spliceosome before step I and detected on a 15% SDS-polyacrylamide gel (lane 1). The investigation of protein-pre-mRNA contacts using anti-Prp2 antibodies indicated that the proteins UV-crosslinked to the U6 RNA did not bind to the substrate. Instead, a polypeptide of 15 kD was in close proximity to the substrate before step I of splicing (lane 3).

High molecular weight proteins UV-crosslinked to the U6 RNA and co-immunoprecipitated with anti-Prp2 antibodies are shown in Figure 40 (Panel B). Six proteins of an apparent molecular size of >200 kD, ~150 kD, 110 kD, a doublet at 90-100 kD, 80 kD were UV-crosslinked to U6 RNA before step I and co-immunoprecipiated with anti-Prp2 antiserum (lane 3). The >200 kD polypeptide may reflect the binding of Prp8p (280 kD) or Slt22p (246 kD) to the U6 RNA in the spliceosome stalled before step I. No denatured protein was immunoprecipitated with anti-Prp2 antibodies, indicating that under the experimental conditions tested the mutant $Prp2^{LAT}$ does not interact directly with U6 RNA (lane 5).

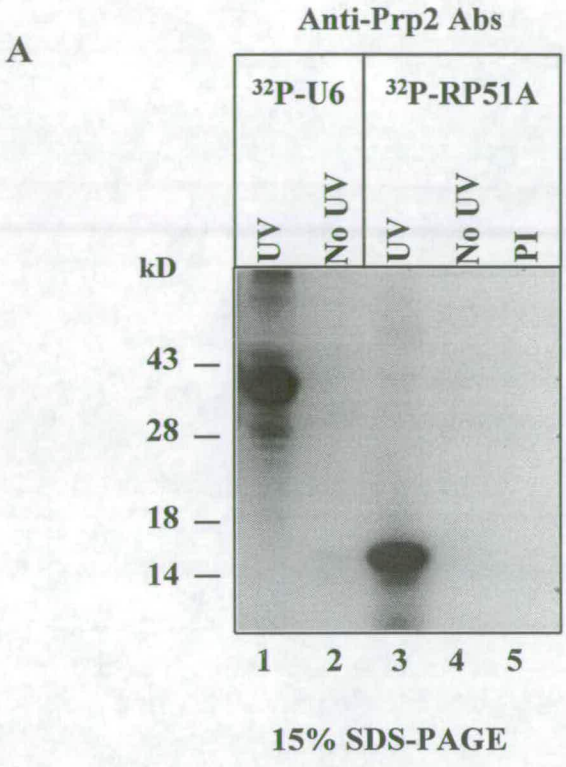
## 9.2 Discussion

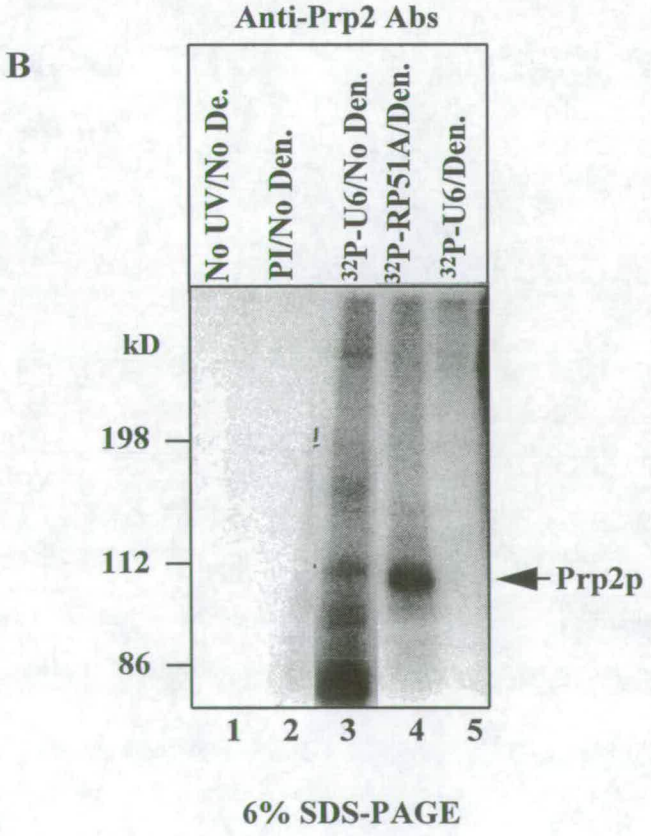
Two proteins (30 kD and 40 kD) specifically interact with the U6 RNA in spliceosomes isolated with all three different antibodies (anti-PAB, -Prp2, -Prp8). Their immunoprecipitation with stalled spliceosomes indicates that they contact U6 RNA at least before step I of splicing. Two additional proteins (23 kD and 25 kD) contacting the U6 RNA in the spliceosome were immunoprecipitated with anti-PAB and anti-Prp8 antisera but were not UV-crosslinked to U6 RNA in spliceosomes stalled before step I and/or immunoprecipitated with anti-Prp2 antisera. This suggests that these two polypeptides bind U6 RNA after the first splicing reaction and/or that compared to anti-Prp2 antibodies, anti-PAB and anti-Prp8 antisera immunoprecipitate spliceosomes more efficiently. These two proteins may contact U6 RNA already in the tri-snRNP since two

**Figure 40. Proteins UV-crosslinked to U6 RNA and immunoprecipitated with anti-Prp2 antibodies.**

**A:** A splicing extract (30 µl) prepared from wild-type SC261 cells (Table 7) was mixed with a splicing extract (20 µl) prepared from S150-2B (pBMPRP2$^{LAT}$) cells (Tables 7 and 9) overproducing the mutant Prp2$^{LAT}$ protein. After depletion of the endogenous U6 snRNA, U6 snRNPs were reconstituted with 6 nM of randomly radiolabelled U6 RNA ($^{32}$P-αATP) (section 2.6.2.3). Wild-type unlabelled substrate RP51A (1 µM final concentration) was added to the reaction samples to induce spliceosome formation. Following U6 snRNP reconstitution and splicing, samples were UV-irradiated and proteins were immunoprecipitated under nondenaturing conditions with anti-Prp2 antisera (lane 1). As a control samples were not UV-irradiated in lane 2. In splicing extract (50 µl) prepared from SC261 cells, a highly radiolabelled substrate RP51A ($^{32}$P-αUTP) (section 2.6.2.1) was added and splicing reaction was carried out. After UV-irradiation, protein samples were immunoprecipitated with anti-Prp2 antibodies (lane 3). As controls sample was not UV-crosslinked (lanes 2 and 4) or proteins were incubated with preimmune serum (lane 5). After washes (150 mM NaCl), proteins were resolved on a 15% SDS-PAGE gel.

**B:** A splicing extract (30 µl) prepared from wild-type SC261 cells (Table 7) was mixed with a splicing extract (20 µl) prepared from S150-2B (pBMPRP2$^{LAT}$) cells (Tables 7 and 9) overproducing the mutant Prp2$^{LAT}$ protein. After depletion of the endogenous U6 snRNA, U6 snRNPs were reconstituted with 6 nM of randomly radiolabelled U6 RNA ($^{32}$P-αATP) (section 2.6.2.3). Wild-type unlabelled substrate RP51A (1 µM final concentration) was added to the reaction samples to induce spliceosome formation. Following U6 snRNP reconstitution and splicing, samples were UV-irradiated. Proteins were immunoprecipitated under nondenaturing conditions (lane 3) or after denaturation (lane 5) with anti-Prp2 antisera. As controls, samples were not UV-irradiated (lane 1), or protein complexes were incubated with the preimmune serum (lane 2). In splicing extract (50 µl) prepared from SC261 cells, a highly radiolabelled substrate RP51A ($^{32}$P-αUTP) (section 2.6.2.1) was addded and splicing reaction was carried out. After UV-irradiation, protein were denatured and immunoprecipitated with anti-Prp2 antibodies (lane 4). Proteins were washed (150 mM NaCl), then resolved on a 6% SDS-PAGE gel.


A
Anti-Prp2 Abs
32P-U6
32P-RP51A
UV
No UV
UV
No UV
PI
kD
43
28
18
14
1
2
3
4
5
15% SDS-PAGE


B
Anti-Prp2 Abs
No UV/No De.
PI/No Den.
32P-U6/No Den.
32P-RP51A/Den.
32P-U6/Den.
kD
198
112
86
Prp2p
1
2
3
4
5
6% SDS-PAGE

proteins of a similar molecular size (23 kD and 25 kD) were immunoprecipitated with anti-Prp8 antisera in the absence of substrate RNA or when a defective RP51A-5'0 pre-mRNA was added in the reaction. It would be interesting to identify the 23 kD and the 25 kD proteins since they possess a molecular weight similar to Uss1p (21 kD).

In metazoans, splicing and cleavage/polyadenylation can be mutually stimulating *in vivo* and *in vitro* (Niwa *et al.*, 1990; Niwa and Berget, 1991). The presence of functional signal for cleavage and polyadenylation can stimulate the removal of the most 3' intron in a pre-mRNA. This involves an interaction between splicing and the 3' end formation machineries (Gunderson *et al.*, 1997; for review see Colgan and Manley, 1997). Although in yeast similar links have not been reported, the binding of PABp to the U6 RNA may reflect an interaction between the polyadenylation and the splicing machineries, however this contact might be an artefact due to the *in vitro* system used.

The investigation of the binding of high molecular weight proteins to the U6 RNA in spliceosomes co-immunoprecipitated with anti-Prp2 antibodies, suggests that $Prp2^{LAT}$ does not directly contact the U6 RNA. This lack of binding may reflect a detrimental effect of the mutation in the putative RNA unwinding motif of Prp2p. Other polypeptides of an apparent molecular weight of 110 kD, ~150 kD, 90-100 kD and 80 kD were also UV-crosslinked to the U6 RNA and co-immunoprecipitated with anti-Prp2 antibodies under nondenaturing conditions. The UV-crosslinking of these proteins to the U6 RNA was difficult to reproduce suggesting that their covalent binding to the U6 RNA in the spliceosome was inefficient and/or the immunoprecipitation of the spliceosomes with anti-Prp2 antibodies was not proficient enough.

## 9.3 Attempts to Investigate of the Presence of Uss1p in the Spliceosome

Proteins interacting with Uss1p are being studied by A. Mayes in the laboratory with the yeast two-hybrid approach (Fromont-Racine *et al.*, 1997; A. Mayes *et al.*, 1997, 1998). Two hybrid screens were carried out using the Gal4-Uss1 fusion protein as a bait. Among many interacting proteins detected, some U2 snRNP-associated proteins, suggesting that Uss1p might be present in the prespliceosome and/or in the spliceosome (A. Mayes, 1997). The UV-crosslinking analysis of proteins interacting with U6 RNA in the spliceosome using anti-Prp8 or anti-PAB antisera allowed us to detect four polypeptides (23 kD, 25 kD, 30 kD and 40 kD). As the 23 kD and the 25 kD proteins possesses a molecular weight similar to Uss1p (21 kD), the presence of this splicing factor in the spliceosome was investigated following several experimental strategies:

***Immunoprecipitation of stalled splicesomes.*** Splicing extracts containing spliceosome complexes stalled before step I were immunoprecipitated with anti-Prp2 antiserum. Proteins were progressively eluted from the complex by gradually increasing the salt concentration of the washing buffer (NTN containing 300 mM-1000 mM NaCl, section 2.8), and precipitated with acetone. Proteins were resolved on an SDS-polyacrylamide gel, and the presence of Uss1p, Prp2p and Prp8p was assayed by Western blotting analysis using the respective antibodies. Only a very weak signal was detected with anti-Prp2 antiserum. The detection of proteins by silver staining indicated that actually few polypeptides had been recovered (data not shown).

***Gel fractionation and elution of splicing complexes.*** After incubation of a $Prp2^{LAT}$-containing splicing extract with a radiolabelled substrate (section 2.4.4.1), splicing complexes (stalled before step I) were resolved on a native gel system (section 2.6.2.2). Complexes were detected by autoradiography and gel slides containing complex I (immature spliceosome containing $Prp2^{LAT}$) or complex III (prespliceosome) were cut out. Gel slices were transferred to dialysis tubes and proteins electoeluted (section 2.9.2). Proteins were either precipitated with acetone or concentrated with a Centricon

centrifugal concentrator (Amicon) according to the manufacturers' instructions. Proteins were resolved on an SDS-PAGE gel and the presence of Uss1p in the splicing complexes was investigated by Western blotting using anti-Uss1 antisera. A faint Uss1p signal was observed in complex III whereas a much stronger band was detected in complex I (data not shown). Despite of the several attempts, we have been unable to repeat this experiment. The native gel system used (section 2.6.2.2) was unreliable and good resolution was difficult to reproduce. Resolution was largely dependent on the quality of the total yeast RNA in the Q buffer required to resolve the spliceosome complex on the native gel system used (section 2.6.2.2). For an unknown reason the best resolution was obtained when total yeast RNA was prepared from temperature-sensitive splicing mutants grown under non-permissive temperature conditions. Therefore the presence of Uss1p in the spliceosome remains unknown.

# CHAPTER TEN

## Final Discussion

Nine proteins are associated with the yeast U6 snRNA: Uss1p (Cooper *et al.*, 1995), Uss2p (Séraphin, 1995), Prp24p (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995), Spb8p, Uss3p, Uss4p, Uss5p, Uss6p and Uss7p (Fromont-Racine *et al.*, 1997; Mayes *et al.*, 1998). Chemical footprinting analysis of the U6 snRNA in the U6 snRNP indicated that only the central domain (in particular nucleotides A40-C43, G50 and G60) and the lower part of the 3' stem-loop (in particular nucleotides G81-A83) of the U6 snRNA are in contact with proteins (Jandrositz and Guthrie, 1995). Further investigations demonstrated that Prp24p contacts the central domain (nucleotides A40-C43) of the U6 snRNA (Jandrositz and Guthrie, 1995; Ghetti *et al.*, 1995).

The work presented here has extended our knowledge on U6 snRNA-protein interactions. Among the U6 snRNP-associated proteins tested Prp24p, Uss1p, Uss3p, Uss4p, Uss5p, Uss6p and Spb8p bind directly to the U6 snRNA. UV-crosslinking analysis demonstrated that at least Uss1p, Uss4p and Uss5p interact with the U6 snRNA in the U6 snRNP, the U4/U6 snRNP and in the tri-snRNP. Also two additional proteins (10 kD and 70-75 kD), as yet unidentified, contact the U6 snRNA in these particles, while other polypeptides (~40 kD, ~85 kD, >210 kD and Prp8p) interact with the U6 snRNA in the tri-snRNP only.

Very little difference, if any, was observed between the proteins bound to the U6 snRNA in the U6 snRNP and in the U4/U6 snRNP. Two explanations are proposed:
1- Prp24p is a U6 snRNP-associated protein which functions at different steps during snRNP assembly. It promotes the annealing of the U6 snRNA and the U4 snRNA to form the U4/U6 snRNP. During this step, Prp24p interacts transiently with these snRNAs in a Prp24p-U4/U6 snRNAs intermediate complex. In this complex Prp24p contacts the U4/U6 snRNAs stem II region, which shows a partial distortion/unwinding. From this observation it has been proposed that Prp24p has to leave this intermediate complex to allow complete U4/U6 snRNAs association. Prp24p may also be involved in the destabilisation of U4/U6 snRNAs as they interact with the spliceosome (Shannon and Guthrie, 1991; Jandrositz and Guthrie, 1995; Ghetti *et al.*,

1995). Recently Prp24p was shown to recycle, in an ATP-dependent manner, the U4/U6 snRNP from U4 and U6 snRNPs released after splicing (Raghunathan and Guthrie, 1998). Although the interaction of Prp24p in the Prp24p-U4-U6 snRNAs complex is transient, the UV-irradiation may have fixed Prp24p with these snRNAs. Consequently, even if the anti-Prp24 antibodies used in this work associate with only traces of U4 snRNA and probably very little of U4/U6 snRNP, following UV-crosslinking, anti-Prp24 antibodies may also precipitate the Prp24p-U4-U6 snRNAs complex and the associated proteins. Thus little difference would have been observed between the proteins contacting the U6 snRNA and isolated with the anti-Prp24 or with anti-Uss1 antibodies, since both antibodies co-immunoprecipitate the U6 and the U4/U6 snRNPs.

2- Despite the transient association of Prp24p with U4/U6 snRNAs, following UV-crosslinking, anti-Prp24 antibodies co-immunoprecipitate mainly the U6 snRNP particle. Thus the little difference observed between proteins contacting the U6 snRNA in the U6 snRNP and in the U4/U6 snRNP indicates that these polypeptides are in close proximity to the U6 snRNA in these two particles. The U6 snRNA undergoes dramatic conformational changes in order to interact with the U4 snRNA in the U4/U6 snRNP. Despite these changes the U6 snRNP associated proteins remain in contact with the U6 snRNA. This suggests that either they accomodate the U6 snRNA conformational changes, or they contact the 3' tail of the U6 snRNA since it is the only region of this molecule which remains single stranded in the U6 snRNP and in the U4/U6 snRNP. Interestingly, the removal of the last 18 nucleotides of the U6 snRNA dramatically affects the binding of proteins to the U6 snRNA in the U6 snRNP and in the U4/U6 snRNP. This suggests that proteins contact the 3' tail of the U6 snRNA or they require this 3' tail to contact the U6 snRNA elsewhere.

Among the U6 snRNP-specific proteins, Uss1p, Uss4p and Uss5p bind to the U6 snRNA in the U6 snRNP, the U4/U6 snRNP and in the U4/U6.U5 tri-snRNP, suggesting an important role for them to play. Several hypotheses can be proposed:

- U6 snRNA undergoes dramatic conformational changes during snRNP assembly and spliceosome formation. Uss1p, Uss4p and Uss5p may function as chaperones and stabilize the U6 snRNA and/or facilitate its conformational changes.
- In yeast, the U6 snRNA is in large excess over the U4 snRNA for no obvious reasons, thus only a fraction of the former associates with the latter to produce the U4/U6 snRNP. It is unknown whether all the U6 snRNA is involved in splicing or if this molecule is required for other cellular function(s). The binding of Uss1p, Uss4p and Uss5p to the U6 snRNA may commit this molecule to splicing and may also be required for the recognition of the U6 snRNA by the other components of the splicing machinery.
- Uss1p, Uss4p and Uss5p share homology with the cannonical Sm proteins. The Sm proteins are required for the hypermethylation of the (7-mG) cap structure of the U1, U2, U4 and U5 snRNAs and also for their nuclear import thus participating at snRNP biogenesis. Although it is unknown whether the yeast U6 snRNA leaves the nucleus, the Sm-like U6 snRNP-specific proteins might be involved in the transport of the U6 snRNA either from the cytoplasm to the nucleus or within the nucleus from one compartment to an other where snRNP biogenesis and/or snRNP assembly occurs.

In addition to Uss1p, Uss4p, Uss5p, a 10 kD and a 70-75 kD proteins, at least five additional proteins (43 kD, 65-80 kD, 85 kD, >>210 kD and Prp8p) contact the U6 snRNA in the tri-snRNP only. The conformation of the U4, U6 and the U5 snRNAs in the tri-snRNP is unknown, however twice as many proteins contact the U6 snRNA in this particle, suggesting that its conformation has changed. It is possible that the U4/U6 snRNA stems I and II have been partially opened-up or distorted, or that the 5' U6 snRNA hairpin structure has been unwound. These conformational changes may have permitted the contact of other proteins with the U6 snRNA. Interestingly, among the proteins bind the U6 snRNA in the tri-snRNP, Prp8p was detected which has been hypothesised to stabilize RNA-RNA interactions between the U5 snRNA and the pre-mRNA (Beggs *et al.*, 1995; Teigelkamp *et al.*, 1995, 1995b; Dix *et al.*, 1998). The >>210 kD protein possesses a size similar to Slt22p (246 kD) which is a U5 snRNP and tri-snRNP associated splicing factor. Slt22p is a member of the DEXH-box protein

family and shows *in vitro* ATPase activity, preferentially stimulated by annealed U2/U6 snRNAs (Xu *et al.*, 1996). One can speculate that in the tri-snRNP, Slt22p unwinds intra- or intermolecular RNA-RNA interactions involving the U6 snRNA and the U4 snRNA. Another tri-snRNP associated protein and member of the DEAD-box protein family of putative RNA helicases, Prp28p (66 kD), was proposed to unwind the U4/U6 snRNAs probably with the help of another factor (maybe Prp24p) (Shannon and Guthrie, 1991; Strauss and Guthrie, 1994). Interestingly a protein of a size (>65 kD) similar to Prp28p binds to the U6 snRNA specifically in the tri-snRNP, indicating that this polypeptide could potentially be Prp28p. Therefore one can speculate that in the tri-snRNP the structure of the U6 snRNA is altered, probably by the action of RNA helicases, and that this modified structure is stabilized by the binding of Prp8p, Uss1p, Uss4p and Uss5p. These U6 snRNA conformational changes may be necessary to prime the association of the tri-snRNP with the prespliceosome.

***Future Work.***

Uss1p, Uss4p and Uss5p bind to the U6 snRNA in the U6 snRNP, U4/U6 snRNP and in the tri-snRNP. It would be of great interest to determine whether these polypeptides contact the U6 snRNA in the spliceosome. The investigation by UV-crosslinking at 254 nm of protein-U6 snRNA contacts in the spliceosome indicated that not many proteins (30 kD, 40 kD) bind to the randomly radiolabelled U6 RNA. In the spliceosome the U6 snRNA is involved in numerous intra- or intermolecular helix structures which might limit the efficiency of protein-U6 snRNA UV-crosslinking. This low efficiency can be partially overcome by using a U6 RNA containing 4-thiouridine. As shown in this work, the UV-crosslinking at 365 nm of proteins to a U6 RNA containing a single radiolabelled 4-thiouridine was highly efficient. Therefore a randomly radiolabelled U6 RNA could be transcribed *in vitro* in a reaction containing a mixture of uridine and 4-thiouridine as well as other nucleotides and components required for transcription. Subsequently, U6 snRNPs would be reconstituted and spliceosomes would form by addition of unlabelled pre-mRNA. Following UV-

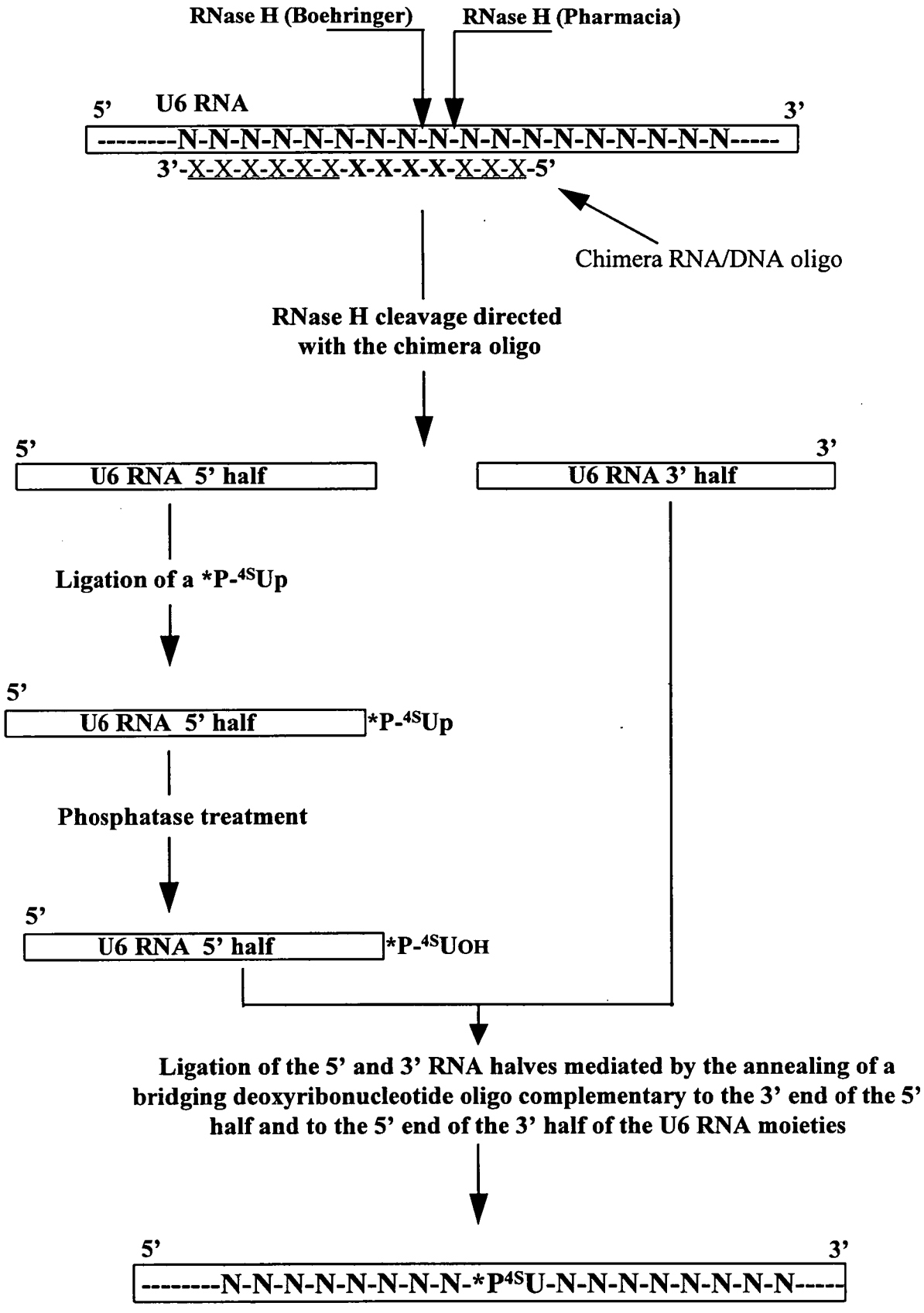
crosslinking at 365 nm, spliceosomes would be co-immunoprecipitated or isolated by gel filtration. After RNase digestion, proteins would be resolved on an SDS-PAGE gel and detected by autoradiography. To investigate the presence of Uss1p, Uss4p and Uss5p, splicing extracts containing as only source of these proteins an HA-fusion would be utilized. After isolation of the spliceosomes and UV-irradiation, proteins would be denatured and immunoprecipitated with anti-HA antibodies. Following resolution of the proteins on an SDS-PAGE gel, those UV-crosslinked to the radiolabelled U6 snRNA would be detected by autoradiography.

Preliminary experiments to map the binding of the proteins to the U6 snRNA showed that several polypeptides (in particular Uss1p and Prp8p) contact the U6 snRNA at position $U_{54}$. It would be of great interest to thoroughly map the binding site of the proteins to the U6 snRNA in the different snRNPs. A photoactivatable 4-thiouridine could be inserted at specific positions of the U6 snRNA following the method described in Figure 41. This method possesses several advantages over that used in this work (Chapter 8 and section 2.4.4.5). Indeed, the 4-thiouridine can be incorporated at any position in the molecule and it does not require the use of 4-thioUpG which is no longer commercially available. As the 5' RNA moiety is generated by RNase H cleavage of the full length U6 RNA, its 3' end does no harbour any non-templated extra nucleotides. Thus, in theory, the ligation efficiency of the 5' and the 3' RNA halves should be higher than that obtained with the method applied in this work.

The binding of Spb8p, Uss3p and Uss6p to the U6 snRNA requires further investigations. Gal4-HA fusion of Spb8p and Uss6p could be generated and the binding of these proteins to the U6 snRNA analysed in the snRNPs. As the Gal4-HA-Uss3p binds weakly to the U6 snRNA, another fusion of this protein could be generated and used to investigate the binding of Uss3p to the U6 snRNA. This might determine whether the fusion of the Gal4-HA to Uss3p has a detrimental effect on its binding to the U6 RNA.

**Figure 41. Site specific incorporation of a $^{32}$P-4 thiouridine.**

This protocol is adapted from that described by Yu and Steitz (1997). Full length U6 snRNA transcribed *in vitro* with the T7 RNA polymerase is annealed with a (2'-O-Methyl-RNA-DNA-2'-O-Methyl RNA) chimera oligo. This chimera oligo is complementary to a sequence of the U6 snRNA and directs its cleavage by RNase H. RNase H cleaves RNA only in an RNA:DNA heteroduplex, whereas RNA base-paired with a 2'-O-Methyl-RNA is protected from cleavage. Depending on the source of RNase H used, the ribonucleotide base-paired with the first deoxyribonucleotide from the 5' end of the chimera oligo cuts either at its 5' site (RNase H purchased from Boehringer) or at its 3' site (RNase H purchased from Pharmacia). A (*P-$^{4S}$Up), derived from ($^{4S}$UpU) by phosphorylation with ($\gamma$-$^{32}$P-ATP) (Laphma *et al.*, 1997) followed by digestion with RNase A, is ligated at the 3' extremity of the 5' half of the U6 RNA. The 3' phosphate of the ligated product is removed by Calf intestinal alkaline phosphatase treatment to produce a 3'-OH group. Subsequently the two halves of the U6 RNA are aligned by complementarity with a bridging oligodeoxyribonucleotide and ligated with the T4 DNA ligase. The resulting U6 RNA contains at a specific site a radiolabelled 4-thiouridine. Open boxes represent the RNA moities. Underlined letters of the chimera oligo represent the 2'-O-Methyl RNA, bold letters represent the deoxyribonucleotide and the RNase H cleavage positions are shown as arrows.


RNase H (Boehringer)
RNase H (Pharmacia)
5' U6 RNA 3'
--------N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-N-----
3'-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-5'
Chimera RNA/DNA oligo
RNase H cleavage directed with the chimera oligo
5'
U6 RNA 5' half
3'
U6 RNA 3' half
Ligation of a *P-4SUp
5'
U6 RNA 5' half
*P-4SUp
Phosphatase treatment
5'
U6 RNA 5' half
*P-4SUOH
Ligation of the 5' and 3' RNA halves mediated by the annealing of a bridging deoxyribonucleotide oligo complementary to the 3' end of the 5' half and to the 5' end of the 3' half of the U6 RNA moieties
5'
3'
--------N-N-N-N-N-N-N-N-*P4SU-N-N-N-N-N-N-N-N-----
Site specifically radiolabelled U6 RNA containing a 4-thioU

In summary, this work describes the first characterisation of the protein contacts made with the U6 snRNA in the snRNP particles and in the spliceosome. Some of these polypeptides appear to accomodate the conformational changes of the U6 snRNA and remain in close proximity to this molecule in the U6, U4/U6 snRNP and in the tri-snRNP, whereas other proteins contact the U6 snRNA only in one type of snRNP. From these observations one can speculate that some of the polypeptides are most likely to maintain the stability of the U6 snRNA while other promote its conformational changes. Interestingly, the U6 snRNP-associated proteins (at least Uss1p, Uss4p and Uss5p) bind to the U6 snRNA in the U6 snRNP, the U4/U6 snRNP and in the tri-snRNP. They possess significant homology with the cannonical Sm proteins, in particular they share the Sm motif. It would be of particular interest to investigate whether by analogy with the core proteins, the U6 snRNP-associated Sm-like proteins are involved in the biogenesis of the U6 snRNP. Do they contact the U6 snRNA in the spliceosome? Are they required for splicing? The U6 snRNA does not contain any cannonical Sm binding site, therefore the mapping of the binding site of the Sm-like proteins onto the U6 snRNA would permit the characterisation of an Sm-like binding site.

# REFERENCES


Abovich, N., Legrain, P. and Rosbash, M. (1990) *Mol. Cell. Biol.*, **10**, 6417-6425.

Abovich, N., Liao, X.L. and Rosbash, M. (1994) *Genes Dev.*, **8**, 843-854.

Abovich, N. and Rosbash, M. (1997) *Cell*, **89**, 403-412.

Aebi, M., Hornig, H., Padgett, R.A., Reiser, J. and Weissmann, C. (1986) *Cell*, **47**, 555-565.

Aebi, M., Hornig, H, and Weismann, C. (1987) *Cell*, **50**, 237-246.

Achsel, T., Teigelkamp, S. Mundt,C., Cronshagen, U., Göthel, S.F., Marahiel, M.A. and Lührmann, R. (1997) *RNA'97*, the second annual meeting of the RNA society.

Ansari, A. and Schwer, B. (1995) *EMBO J.*, **14**, 4001-4009.

Anthony, J.G., Weidenhammer, E.M. and Woolford, J.L.JR. (1997) *RNA*, **3**, 143-1152.

Arenas, J.E. and Abelson, J.N. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 6771-6775.

Arenas, J.E. and Abelson, J.N. (1998) *Proc. Natl. Acad. Sci. USA,* **94**, 11798-11802.

Arenas, J. and Hurwitz, J. (1987) *Biol. Chem.,* **25**, 4274-4279.

Ares, M. and Weiner, B. (1995) *Prog. Nucleic Acid Res. Mol. Biol.*, **50**, 131-159.

Arning, S., Grüter, P., Bilbe, G. and Krämer, A. (1996) *RNA*, **2**, 794-810.

Ast, G. and Weiner, A.M. (1996) *Science*, **272**, 881-884.

Ast, G. and Weiner, A.M. (1997) *RNA*, **3**, 371-381.

Ayadi, L., Miller, M. and Banroques, J. (1997) *RNA*, **3**, 197-209.

Bach, M., Winkelmann, G. and Lührmann, R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 6038-6042.

Banroques, J., Abelson, J. (1989) *Mol. Cell. Biol.*, **9**, 3710-3719.

Barabino, S.M., Blencowe, B.J., Ryder, U., Sproat, B.S. and Lamond, A.I. (1990) *Cell,* **63**, 293-302.

Been, M.D. and Perrotta, A.T. (1991) *Science*, **252**, 434-437.

Beggs, J.D., Van den Berg, J., Ooyen, A. and Weissmann, C. (1980) *Nature*, **283**, 835-840.

Beggs, J.D. and M. Plumpton (1992) *Nucleic Acids and Mol. Biol.*, **6**, 187-202.

Beggs, J.D. (1993) *Mol. Biol. Reports*, **18**, 99-103.

Beggs, J.D. (1995) Pre-mRNA processing, A. Lamond (ed), Springer-Verlag Berlin Heidelberg. pp79-95.

Beggs, J.D, Teigelkamp, S. and Newman, A.J. (1995) *J. Cell Science*, **19**, 101-105.

Behrens, S-E., Tyc, K., Kastner, B., Reichelt, J. and Lührmann, R. (1993) *Mol. Cell. Biol.*, **13**, 307-319.

Bennett, M., Michaud, S., Kingston, J. and Reed, R. (1992) *Genes Dev.*, **6**, 1986-2000.

Bennett, M. Reed, R. (1993) *Science*, **262**, 105-108.

Berglund, J.A., Chua, K., Abovich, N., Reed, R. and Rosbash, M. (1997) *Cell*, **89**, 781-787.

Birney, E., Kumar, S. and Krainer, A.R. (1993) *Nucleic Acids Res.*, **21**, 5803-5816.

Blanton, S., Srinivasan, A. and Rymond, B. (1992) *Mol. Cell. Biol.*, **12**, 3939-3947.

Blencowe, B.J., Sproat, and Ryder, U. (1989) *Cell*, **59**, 531-539.

Boelens, W.C., Scherly, D., Jansen, E.J.R., Kolen, K., Mattaj, I.W. and Van Venrooij, W.J. (1991) *Nucleic Acids Res.*, **19**, 4611-4618.

Boelens, W.C., Palacios, I. and Mattaj, I.W. (1995) *RNA*, **1**, 273-283.

Boeck, R., Lapeyre, B., Brown, C.E. and Sachs, A.B. (1998) *Mol. Cell. Biol.*, in press.

Bordonné, R., Banroques, J., Abelson, J. and Guthrie, C. (1990) *Genes Dev.*, **4**, 1185-1195.

Branlant, C., Krol, A., Ebel, J.P., Lazar, E., Haendler, B. and Jacob, M. (1982) *EMBO J.*, **1**, 1259-1265.

Breathnach, R. and Chambon, P. (1981) *Annu. Rev. Biochem.*, **50**, 349-384.

Breathnach, R., Manel, J.L. and Chambon, P. (1977) *Nature*, **270**, 314-319.

Bringmann, P., Renter, R., Rinke, J., Appel, B., Bald, R. and Lührmann, R. (1984) *J. Biol. Chem.*, **258**, 2745-2747.

Brosi, R., Hauri, H-P. and Krämer, A. (1993) *J. Biol. Chem.*, **268**, 17640-17646.

Brosi, R., Gröning, K., Behrens, S-E., Lührmann, R. and Krämer, A. (1993b) *Science*, **262**, 102-105.

Brow, D.A. and Guthrie, C. (1988) *Nature*, **334**, 213-218.

Brown, J.D. and Beggs, J.D. (1992) *EMBO J.*, **11**, 3721-3729.

Brys, A. and Schwer, B. (1996) *RNA*, **2**, 707-717.

Burke, J.M. and RajBhandary, U.L. (1982) *Cell*, **31**, 509-520.

Camasses, A., Bragado-Nilsson, E., Martin, R., Séraphin, B. and Bordonné, R. (1998) *Mol. Cell. Biol.*, **18**, 1956-1966.

Cazenave, C. and Uhlenbeck, C.O. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 6972-6976.

Cech, T.R. (1985) *Cell*, **43**, 713-716.

Cech, T.R. (1993) *The RNA world*, Cold Spring Harbor Laboratory Press, p239-269.

Champion-Arnaud, P. and Reed, R. (1994) *Genes Dev.*, **8**, 1974-1983.

Champion-Arnaud, P., Gonzani, O., Palandjian, L. and Reed, R. (1995) *Mol. Cell. Biol.*, **15**, 5750-5756.

Chanfreau, G., Elela, S.A., Ares, M.Jr. and Guthrie, C. (1997) *Genes Dev.*, **11**, 2741-2751.

Chapman, K.B. and Boeke, J.D. (1991) *Cell*, **65**, 483-492.

Chapon, C. and Legrain, P. (1992) *EMBO J.*, **11**, 3279-3288.

Chen, H.R., Jan, S.P., Tsao, T.Y., Sheu, Y.J., Banroques, .and Cheng, S.C. (1998) *Mol. Cell Biol.* ;**18**, 2196-2204.

Cheng, S-C. and Abelson, J. (1987) *Genes Dev.*, **1**, 1014-1027.

Chiara, M.D., Champion-Arnaud, P., Buvoli, M., Nadal-Ginard, B. and Reed, R. (1993) *Proc. Natl. Acad. Sci. USA*, **91**, 6403-6407.

Chiara, M.D., Champion-Arnaud, P., Buvoli, M., Nadal-Ginard, B., and Reed, R. (1994) *Proc. Natl. Acad. Sci. U S A*, **91**, 6403-6407.

Christopher, D.A. and Hallick, R.B. (1989) *Nucleic Acids Res.*, **17**, 7591-7608.

Chu, F., Maley, G.F., West, D.K., Belfort, M. and Maley, F. (1986) *Cell*, **45**, 157-166.

Chu, J-l. and Elkon, K.B. (1991) *Gene*, **97**, 311-312.

Church, G.M. and Gilbert, G. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 1991-1995.

Ciliberto, G., Dathan, N., Frank, R., Philipson, L. and Mattaj, I.W. (1986) *EMBO J.*, **5**, 2931-2937.

Colgan, D.F. and Manley, J.L. (1997) *Genes Dev.*, **11**, 2755-2766.

Colot, H.V., Stutz, F. and Rosbash, M. (1996) *Genes Dev.*, **10**, 1699-1708.

Company, M., Arenas, J. and Abelson, J. (1991) *Nature*, **349**, 487-493.

Cooper, M. (1995) Ph.D. Thesis, University of Edinburgh.

Cooper, M, Johnston, L.H. and Beggs, J.D. (1995) *EMBO J.*, **14**, 2066-2075.

Copertino, D.W. and Hallick, R.B. (1991) *EMBO J.*, **10**, 433-442.

Copertino, D.W.,Christopher, D.A. and Hallick, R.B. (1991b) *Nucleic Acids Res.*, **19**, 6491-6497.

Copertino, D.W., Shigeoka, S. and Hallick, R.B. (1992) *EMBO J.*, **11**, 5041-5050.

Copertino, D.W. and Hallick, R.B. (1993) *TIBS*, **18**, 467-471.

Crispino, J.D., Blencowe, B.J. and Sharp, P.A. (1994) *Science*, **265**, 1866-1869.

Dahlberg, J.E. and Lund, E. (1988) Birnstiel, M.L., Verlag-Springer, pp 38-70.

Dalbadie-McFarland and Abelson (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4236-

Dalrymple, M.A., Petersen-Bjorn, S.; Friesen, J.D. and Beggs, J.D. (1989) *Cell*, **58**, 811-812.

Das, G., Henning, D. and Reddy, R. (1987) *J. Biol. Chem.*, **262**, 1187-1193.

Datta, B. and Weiner, A.M. (1991) *Nature*, **352**, 821-824.

Datta, B. and Weiner, A.M. (1993) *Mol. Cell. Biol.*, **13**, 5377-5382.

Davies, R.W, Waring, R.B., Ray, J.A., Brown, T.A. and Scazzocchio, C. (1982) *Nature*, **300**, 719-724.

Decatur, W.A., Einvik, C., Johansen, S. and Vogt, V.M. (1995) *EMBO J*, **14**, 4558-4568.

Dix, I., O'Keefe, R., Russel, C., Newman, A. and Beggs, J.D. (1998) *RNA, In press.*

Eperon, I.C., Ireland, D.C., Smith, R.A., Mayeda, A. and Krainer, A.R. (1993) *EMBO J.*, **12**, 3606-3617.

Fabrizio, P., McPheeters, D.S. and Abelson, J. (1989) *Genes Dev.*, **3**, 2137-2150.

Fabrizio, P. and Abelson, J. (1990) *Science,* **250**, 404-409.

Fabrizio, P. and Abelson, J. (1992) *Nucleic Acids Res.*, **20**, 3659-3664.

Fabrizio, P., Esser, S., Kastner, B. and Lührmann, R. (1994) *Science*, **264**, 261-264.

Fabrizio, P., Laggerbauer, B., Lauber, J., Lane, W.S. and Lührmann, R. (1997) *EMBO J.*, **16**, 4092-4106.

Favre, A. (1990) Morrison, H (ed.) *Bioorganic Photochemistry*, **Chap. 5**, 379-425.

Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.*, **132**, 6-13.

Feinberg, A.P. and Vogelstein, B. (1984) *Addendum. Anal. Biochem.*, **137**, 266-267.

Fetzer, S., Lauber, J., Will, C.L and Lührmann, R. (1997) *RNA*, **3**, 334-355.

Field, D.J and Friesen, J.D. (1996) *Genes Dev.*, **10**, 489-501.

Fisher, D.E, Conner, G.E., Reeves, W.H., Wisniewolski, R. and Blobel, G. (1985) *Cell*, **42**, 751-758.

Fisher, U., Wittman-Liebold, B., Lang, K., Kiefhaber, T. and Schmid, F.X. (1989) *Nature*, **337**, 476-478.

Fisher, U. and Lührmann, R. (1990) *Science*, **249**, 786-790.

Fleckner, J., Zhang, M., Valcarcel, J. and Green, M.R. (1997) *Genes Dev.*, **15**, 1864-1872.

Foreman, P.K. and Davis, R. (1994) *Gene*, **144**, 63-68.

Forné, T., Rossi, F., Labourier, E., Antoine, E., Cathala, G., Brunel, C. and Tazi, J. (1995) *J. Biol. Chem.*, **270**, 16476-16481.

Fortner, D.M., Troy, R.G. and Brow, D.A. (1994) *Genes Dev.*, **8**, 221-233.

Fouser, L. A. and Friesen, J.D. (1986) *Cell*, **45**, 81-93.

Frank, D.N., Patterson, B. and Guthrie, C. (1992) *Mol. Cell. Biol.*, **12**, 5197-5205.

Frank, D.N., Roiha, H. and Guthrie, C. (1994) *Mol. Cell. Biol.*, **14**, 2180-2190.

Frendeway, D. and Keller, W. (1985) *Cell*, **42**, 355-367.

Fresco, L.D. and Buratowski, S. (1996) *RNA*, **2**, 584-596.

Fromont-Racine, M., Rain, J-C. and Legrain, P. (1997) *Nature Genetics*, **16**, 277-282.

Fu, X-D. and Maniatis, T. (1990) *Nature*, **343**, 437-441.

Fu, X-D. (1995) *RNA*, **1**, 663-680.

Fuller-Pace, F.V. and Lane, D. (1992) In Eckstein, F. and Lilley, D.M.J. (eds), Nucleic acids and Molecular Biology Volume 6. Springer-Verlag Berlin Heidelberg. pp. 159-173.

Fury, M. and Zieve, G.W. (1996) *Experimental Cell Research*, **228**, 160-163.

Galisson, F. and Legrain, P. (1993) *Nucleic Acids Res.*, **21**, 1555-1562.

Garcia-Blanco, M.A., Anderson, G.J., Beggs, J.D. and Sharp, P.A. (1990) *Proc. Natl. Acad. Sci.* USA, **87**, 3082-3086.

Ge, H. and Manley, J.L (1990) *Cell*, **62**, 25-34.

Gee, S., Wald Krauss, S., Miller, E., Aoyagi, K., Arenas, J. and Conboy, J.G. (1998) *Proc. Natl. Acad. Sci. USA*, **94**, 11803-11807.

Ghetti, A., Company, M. and Abelson, J. (1995) *RNA*, **1**, 132-145.

Gietz, R. D. and Sugino, A. (1988) *Gene*, **74**, 527-535.

Gietz, R. D., St Jean, A., Woods, R.A. and Schiestl, R.H. (1992) *Nucleic Acids Res.*, **21**, 1425.

Goldschmidt-Clermont, M. (1991) *Cell*, **65**, 135-143.

Gott, J.M., Shub, D.A and Belfort, M. (1986) *Cell*, **47**, 81-87.

Gottschalk. A., Tang, J., Puig, O., Salgado, J., Neubauer, G., Colot, H.V., Mann, M., Séraphin, B., Rosbash, M., Lührmann, R. and Fabrizio, P. (1998) *RNA*, **4**, 374-393.

Gozani, O., Patton, J.G. and Reed, R. (1994) *EMBO J.*, **13**, 3356-3367.

Gozani,O., Feld, R. and Reed, R. (1996) *Genes Dev.*, **10**, 233-243.

Greenberg, J.R. (1979) *Nucl. Acids Res.*, **6**, 715-732.

Gunderson, S.I., Vagner, S., Polycarpou-Schwarz, M. and Mattaj, I.W. (1997) *Genes Dev.*, **11**, 761-773.

Guthrie, C. and Patterson, B. (1988) *Annu. Rev. Genet.*, **22**, 387-419.

Hall, I.J. (1991) *Nucleic Acids Res.*, **19**, 3795-3798.

Hall, R.S. and Padgett, R.A. (1994) *J. Mol.Biol.*, **239**, 357-365.

Hall, S.L and Pradgett, R.A. (1996) *Science*, **271**, 1716-1718.

Hallick, R.B., Hong, L., Drager, R.G., Favreau, M.R., Monfort, A., Orsat, B., Spielmann, A. and Stutz, E. (1993) *Nucleic Acids Res.*, **21**, 3537-3544.

Haltiner-Jones, M. and Guthrie, C. (1990) *EMBO J.*, **9**, 2555-2561.

Hamm, J., Datzynkiewicz, E., Tahara, S.M. and Mattaj, I.W. (1990) *Cell*, **62**, 569-577.

Hanahan, D. (1983) *J. Mol. Biol.*, **166**, 557-580.

Harrison, S. (1991) *Nature*, **353**, 715-719.

Hashimoto, C. and Steitz, J.A. (11984) *Nucleic Acids Res.*, **12**, 3283-3293.

Heinrichs, V., Bach, M., Winkelmann, G. and Lührmann, R. (1990) *Science,* **247**, 69-72.

Heinrichs, V., Hackl, W. and Lührmann, R. (1992) *J. Mol. Biol.*, **227**, 15-28.

Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H. and Lührmann, R. (1995) *EMBO J.*, **14**, 2076-2088.

Hernandez, N. and Keller, W. (1983) *Cell*, **35**, 89-99.

Hernandez, N. and Weiner, A.M. (1986) *Cell*, **47**, 249-258.

Hetzer, M., Wurzer, G., Schweyen, R.J and Manfred, W.M. (1997) *Nature*, **386**, 417-420.

Hieter, P. and Sikorski, R.S. (1989) *Genetics*, **122**, 19-27.

Hilleren, P.J., Kao, H-Y. and Siliciano, P.G. (1995) *Mol. Cell. Biol.*, **15**, 6341-6350.

Hodges, P.E. and Beggs, J.D. (1994) *Current Biology*, **4**, 264-267.

Hodges, P.E., Jackson, S.P., Brown, J.D. and Beggs, J.D. (1995) *Yeast*, **11**, 337-342.

Hong, L. and Hallick, R.B. (1994) *Genes & Dev.*, **8**, 1589-1599.

Hong, W., Bennett, M., Xiao, Y., Feld Kramer, R., Wang, C. and Reed, R. (1997) *Nucleic Acids Res.*, **25**, 354-361.

Horowitz, D.S. and Abelson, J. (1993a) *Mol. Cell. Biol.*, **13**, 2959-2970.

Horowitz, D.S. and Abelson, J. (1993b) *Genes Dev.*, **7**, 320-329.

Horowitz, D.S. and Krainer, A. (1997) *Genes Dev.*, **11**, 139-151.

Hu, J., Xu, Y., Schappert, K., Harrington, T., Wang, A., Braga, R., Mogridge, J. and Friesen, J.D. (1994) *Nucleic Acids Res.*, **22**, 1724-1734.

Igel, H., Wells, S., Perriman, R. and Ares, M. Jr. (1998) *RNA*, **4**, 1-10.

Iggo, R.D., Jamieson, D.J., MacNeill, S.A., Southgate, J., McPheat, J. and Lane, D.P. (1991) *Mol. Cell. Biol.*, **11**, 1326-

Izaurralde, E. and Adam, S. (1998) *RNA*, **4**, 351-364.

Jackson, S.P., Lossky, M. and Beggs, J.D. (1988) *Mol. Cell. Biol.*, **8**, 1067-1075.

Jacquier, A. and Rosbash, M. (1986) *Proc. Natl. Acad. Sci. USA.*, **83**, 5835-5839.

Jacquier, A. (1990) *Trends Biochem. Sci.*, **15**, 351-354.

Jacquier, A. and Jacquesson-Breuleux, N. (1991) *J. Mol. Biol.*, **219**, 415-428.

Jamison, S.F., Crow, A. and Garcia-Blanco, M.A. (1992) *Mol. Cell. Biol.*, **12**, 4279-4287.

Jandrositz, A. and Guthrie, C. (1995) *EMBO J.*, **14**, 820-832.

Jarmolowski, A. and Mattaj, I.W. (1993) *EMBO J.*, **12**, 223-232.

Jarrell, K.A., Dietrich, R.C. and Perlamn, P.S. (1988) *Mol. Cell. Biol.*, **8**, 2361-2366.

Jeffreys, A.J. and Flavell, R. A. (1977) *Cell*, **12**, 1097-1108.

Jenkins, K.P, Hong, L. and Hallick, R.B. (1995) *RNA*, **1**, 624-633.

Johansen, S., Elde, M., Vader, A., Haugen, P., Haugli, K and Haugli, F. (1997) *Molecular Microbiology*, **24**, 737-745.

Johnston, M. and Davis, R.W. (1984) *Mol. Cell. Biol.*, **4**, 1440-1448.

Jones, M.H. and Guthrie, C. (1990) *EMBO J.*, **9**, 2555-2561.

Jones, M.H., Frank, D.N. and Guthrie, C. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 9687-9691.

Kambach, C. and Mattaj, I.W. (1992) *J. Cell. Biol.*, **118**, 11-21.

Kandel-Lewis, S. and Séraphin, B. (1993) *Science*, **262**, 2035-2039.

Kao, H-Y. and Siliciano, P.G. (1992) *Nucleic Acids Res.*, **20**, 4009-4013.

Kao, H-Y. and Siliciano, P.G. (1996) *Mol. Cell. Biol.*, **16**, 960-967.

Kao, H-Y. and Rymond, B. (1996) *Mol. Cell. Biol.*, **16**, 960-967.

Keller, E.B. and Noon, W.A. (1985) *Nucleic Acids Res.*, **13**, 4971-4981.

Kenan, D.J., Query, C.C. and Keene, J.D. (1991) *Trends Biochem. Sci.*, **16**, 214-220.

Kidd, D.D. and Ruano, G. (1994) PCR 2: A practical approach. IRL Press, **p**.9.

Kim, S.H., Smith, J., Claude, A. and Lin, R.J. (1992) *EMBO J.*, **11**, 2319-2326.

Kim, S-H. and Lin, R-J. (1996) *Mol. Cell. Biol.*, **16**, 6810-6819.

Kim, C.H., Ryan, D.E., Marciniec, T. and Abelson, J. (1997) *EMBO J.*, **16**, 2119-2129.

King, D.S. and Beggs, J.D (1990) *Nucleic Acids Res.*, **18**, 6559-6564.

Kiss, T., Jakab, G., Antal, M., Pálfi, Z., Hegyi, H., Kis, M. and Solymosy, F. (1988) *Nucleic Acids Res.*, **16**, 5407-5426.

Kleinschmidt, A.N. and Pederson, T. (1987) *Mol. Cell. Biol.*, **7**, 3131-3137.

Kleinschmidt, A.N., Patton, J.R. and Pederson, T. (1989) *Nucleic Acids Res.*, **17**, 4817-4828.

Koch, J.L., Boulanger, S.C., Dibhajj, S.D, Hebbar, S.K and Perlman, P.S. (1992) *Mol. Cell. Biol.*, *12*, 1950-1958.

Koerner, T.J., Hill, J.E., Myers, A.M. and Tzagoloff, A. (1990) *Methods in Enz.*, **194**, 477-490.

Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R. and Garcia-Blanco, M.A. (1994) *Nature*, **368**, 119-124.

Kolossova, I. and Padgett, R.A. (1997) *RNA*, **3**, 227-233.

Konarska, M.M., Grabowski, P.J., Padgett, R.A. and Sharp, P.A. (1985) *Nature*, **313**, 552-557.

Konarska, M.M. and Sharp, P.A. (1986) *Cell*, **46**, 845-855.

Konarska, M.M. and Sharp, P.A. (1987) *Cell*, **49**, 763-774.

Konforti, B.B. and Konarska, M.M. (1994) *Genes Dev.*, **8**, 1962-1973.

Krainer, A.R., Maniatis, T., Ruskin, B. and Green, M.R. (1984) *Cell*, **36**, 993-1005.

Krainer, A.R., Conway, G.C. and Kozak, D. (1990) *Cell*, **62**, 35-42.

Krämer, A., Legrain, P., Mulhauser, F., Gröning, K., Brosi, R. and Bilbe, G. (1994) *Nucleic Acids Res.*, **22**, 5223-5228.

Krämer, A. (1995) In Lamond, A.I. (ed.), Pre-mRNA Processing, R.G. Landes Company, Springer-Verlag, pp. 35-64.

Kunkel, TA., Roberts, JD. and Zamour, RA. (1987) *Methods in Enz.*, **154**, 367-382.

Laemmli, U.K. (1970) *Nature*, **277**, 680-685.

Laggerbauer, B., Lauber, J. and Lührmann, R. (1996) *Nucleic Acids Res.*, **24**, 868-875.

Lamb, J.R.; Tugendreich, S. and Heiter, P. (1995) *Trends Biochem. Sci.*, **20**, 257-259.

Lambowitz, A. M. and Belfort, M. (1993) *Annu. Rev. Biochem.*, **62**, 587-622.

Lambright, D.G., Sondek, J., Bohm, A., Skiba, N.P., Hamm, H.E., Sigler, P.B. (1996) *Nature*, **376**, 311-319.

Lamm, G.M., Blencowe, B.J. and Sproat, B.S. (1991) *Nucleic Acids Res.*, **19**, 3193-3198.

Lamond, A.I., Konarska, M.M. and Sharp, P.A. (1987) *Genes Dev.*, **1**, 532-543.

Lamond, A.I., Konarska, M.M. and Grabowski, P.A. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 411-415.

Lapham, J, Yu, Y-T., Shu, M-D., Steitz, J.A. and Crothers, D.M. (1997) *RNA*, **3**, 950-951.

Last, R.L., Maddock, J.R. and Woolford, J.L. (1987) *Genetics*, **117**, 619-631.

Lauber, J., Fabrizio, P., Teigelkamp, S., Lane, W.S., Hartmann, E. and Lührmann, R. (1996) *EMBO J.*, **15**, 4001-4015.

Lauber, J., Plessel, G., Prehn, S., Will, C.L., Fabrizio, P., Grönig, K., Lane, W.S. and Lührmann, R. (1997) *RNA*, **3**, 926-941.

Lazowska, J., Claisse, M., Gargouri, A., Zotylak, Z., Spyridakis, A. and Slonimski, P.P. (1989) *J. Mol. Biol.*, **205**, 275-290.

Legrain, P., Séraphin, B. and Rosbash, M. (1988) *Mol. Cell. Biol.*, **8**, 3755-3760.

Legrain, P., Chapon, C. and Galisson, F. (1993) *Genes Dev.*, **7**, 1390-1399.

Legrain, P., Chapon, C. (1993b) *Science*, **262**, 108-110.

Lehmeier, T., Raker, V., Hermann, H. and Lührmann, R. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 12317-12321.

Lesser, C.F. and Guthrie, C. (1993) *Science*, **262**, 1982-1988.

Lewis, J.D., Izaurralde, E., Jarmolowski, A., McGuigan, C. and Mattaj, I.W. (1996a) *Genes Dev.*, **10**, 1683-1698.

Lewis, J.D., Görlich, D. and Mattaj, I.W. (1996b) *Nucleic Acids Res.*, **24**, 3332-3336.

Li, Z. and Brow, D.A. (1993) *Nucleic Acids Res.*, **21**, 4645-4646.

Liautard, J.P., Sri-Widada, J., Brunel, C. and Jeanteur, P. (1982) *J. Mol. Biol.*, **162**, 623-643.

Liao, X.C., Colot, H.V., Wang, Y. and Rosbash, M. (1992) *Nucleic Acids Res.*, **20**, 4237-4245.

Liao, X.C., Tang, J. and Rosbash, M. (1993) *Genes Dev.*, **7**, 419-428.

Lin, R.-J., Newman, A.J., Cheng, S.-C. and Abelson, J. (1985) *J. Biol. Chem.*, **260**, 14780-14792.

Lin, R.-J., Lustig, A.J. and Abelson, J. (1987) *Genes Dev.*, **1**, 7-18.

Lin, J. and Rossi, J.J. (1996) *RNA*, **2**, 835-848.

Lockhart, S.R. and Rymond, B. (1994) *Mol. Cell. Biol.*, **14**, 3623-3633.

Lossky, M., Anderson, G.J., Jackson, S. and Beggs, J.D. (1987) *Cell*, **51**, 1019-1026.

Lührmann, R., Kastner. B. and Bach, M. (1990) *Biochim. Biophys. Acta.*, **1087**, 265-292.

Lund, E. and Dahlberg, J.E. (1992) *Science*, **255**, 327-330.

Maddock, J.R., Roy, J., and Woolford, J.L. Jr. (1996) *Nucleic Acids Res.*, **6**, 1037-1044.

Madhani, H.D., Bordonné, R. and Guthrie, C. (1990) *Genes Dev.*, **4**, 2264-2277.

Madhani, H.D. and Guthrie, C. (1992) *Cell*, **71**, 803-817.

Madhani, H.D. and Guthrie, C. (1994a) *Genes Dev.*, **8**, 1071-1086.

Madhani, H.D. and Guthrie, C. (1994b) *Genetics*, **137**, 677-687.

Madhani, H.D. and Guthrie, C. (1994c) Annu. Rev. *Genet.*, **28**, 1-26.

Madore, S.J., Wieden, E.D. and Pederson, T. (1984a) *J. Cell. Biol.*, **98**, 188-192.

Madore, S.J., Wieden, E.D., Kunkel, G.R. and Pederson, T. (1984b) *J. Cell. Biol*, **99**, 1140-1144.

Maier, U.G., Rensing, S.A., Igloi, G.L. and Maez, M. (1995) *Mol. Gen. Genet.*, **246**, 128-131.

Manley, J.L., Tacke, R. (1996) *Genes Dev.*, **10**, 1569-1579.

Mattaj, I.W. and De Robertis, E.M. (1985) *Cell*, **40**, 111-118.

Mattaj, I.W. (1986) *Cell*, **46**, 905-911.

Mattaj, I.W., Habets, W.J., van Venrooij, W.J. (1986) *EMBO J.*, **5**, 997-1002.

Mattaj, I.W. (1988) (Birnstiel M.L ed), Berlin, Springer-Verlag.

Mayes, A. E., Fromont-Racine, M., Legrain, P. and Beggs, J.D. (1997) RNA'97, The second annual meeting of the RNA society (Banff Centre).

Mayes, A. E., Legrain, P. and Beggs, J.D. (1998) In preparation.

McAllister, G., Roby, S.A., Amara, S.G. and Lerner, M.R. (1989) *EMBO J.*, **8**, 1177-1181.

McLean, M.R. and Rymond, B.C. (1998) *Mol. Cell. Biol.* ;*18*, 353-360.

McPheeters, D.S. (1996) *RNA*, **2**, 1110-1123.

McPheeters, D.S., Fabrizio, P. and Abelson, J. (1989) *Genes Dev.*, **3**, 2124-2136.

McPheeters, D.S. and Abelson, J. (1992) *Cell*, **71**, 819-831.

Mermound, J.E., Calvio, C. and Lamond, A.I. (1994) *Adv. Prot. Phosphatases*, **8**, 99-118.

Messing, J., Gronenborn, B., Muller-Hill, B. and Hofschneider, P.H. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 3642-3646.

Michaud, S. and Reed, R. (1991) *Genes Dev.*, **5**, 2534-2546.

Michel, F., Hanna, M., Green, R., Bartel, D.P. and Szostak, J.W. (1989) *Nature*, **342**, 391-395.

Michel, F., Umesono, K and Ozeki, H. (1989b) *Gene*, **82**, 5-30.

Michel, F. and Westhof, E. (1990) *J. Mol. Biol.*, **216**, 585-610.

Milligan, J.F. and Uhlenberck, O.C. (1989) *Methods in Enz.*, **180**, 51-62.

Montzka Wassarman, K. and Steitz, J.A. (1992) *Mol. Cell. Biol.*, **12**, 1276-1285.

Moore, M.J., Query, C. and Sharp, P.A. (1993) *The RNA world*, Cold Spring Harbor Laboratory Press, **Chapter 13**, pp 303-357.

Mount, S.M. (1982) *Nucleic Acids Res.*, **10**, 459-472.

Mueller, M.W., Allmaier, M., Eskes, R. and Schweyen, R.J. (1993) *Nature*, **366**, 174-176.

Nakagawa, T. and Ogawa, H. (1997) *Genes Cells*, **2**, 65-79.

Neer, E.J., Schmidt, C.J., Nambudripad, R., Smith, T.F. (1994) *Nature*, **371**, 297-300.

Neer, E.J. (1995) *Cell*, *80*, 249-257.

Nelissen, R.L.H., Heinrichs, V., Habets, W.J., Simons, F., Lührmann, R. and Van Venrooij, W.J. (1991) *Nucleic Acid Res.*, **19**, 449-454.

Nelissen, R.L.H., Will, C.L., Van Venrooij, W.J. and Lührmann, R. (1994) *EMBO J.*, **13**, 4113-4125.

Neubauer, G., Gottschalk, A., fabrizio, P., Séraphin, B., Lührmann, R. and Mann, M. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 385-390.

Neuman de Vegvar, H.E. and Dahlberg, J.E. (1986) *Cell*, **47**, 259-266.

Neuman de Vegvar, H.E. and Dahlberg, J.E. (1990) *Mol. Cell. Biol.*, **10**, 3365-3375.

Newman, A.J., Lin, R.-J., Cheng, S.-C., and Abelson, J. (1985) *Cell*, **42**, 335-344.

Newman, A.J. and Norman, C. (1991) *Cell*, **65**, 115-123.

Newman, A.J. and Norman, C. (1992) *Cell*, **68**, 743-754.

Newman, A.J.(1994) *Curr. Biol.*, **4**, 462-464.

Newman, A.J., Teigelkamp, S. and Beggs, J.D. (1995) *RNA*, **1**, 968-980.

Newman, A.J. (1997) *EMBO J.*, **16**, 5797-5800.

Nilsen, T.W. (1994) *Cell*, **78**, 1-4.

Niwa, M., Rose, S.D. and Berget, S.M. (1990) *Genes Dev.*, **4**, 1552-1559.

Niwa, M. and Berget, S.M. (1991) *Genes Dev.*, **5**, 2086-2095.

Noble, S.M. and Guthrie, C. (1996a) *Genetics*, **143**, 67-80.

Noble, S.M. and Guthrie, C. (1996b) *EMBO J.*, **15**, 4368-4379.

O'Day, C.L., Dalbadie-McFarland, G. and Abelson, J. (1996) *J. Biol. Chem.*, **271**, 33261-33267.

Ogawa, H.; Johzuka, K.; Nakagawa, T.; Leem, S.H. and Hagihara, A.H. (1995) *Adv Biophys*, **31**, 67-76.

O'Keefe, R.T., Norman, C. and Newman, A.J. (1996) *Cell*, **86**, 679-689.

O'Keefe, R.T. and Newman, A.J. (1998) *EMBO J.*, **17**, 565-574.

Ordway, R.W., Pallanck, L. and Ganetzky, B. (1995) *Trends Biochem. Sci.*, **19**, 530-531.

Padgett, R.A., Garbowski, P.J., Konarska, M.M., Seiler, S. and Sharp, P.A. (1986) *Annu. Rev. Biochem.*, **55**, 1119-1150.

Pante N., Jarmolowski A., Izaurralde E., Sauder U., Baschong W. and Mattaj I.W. (1997) *RNA*, **3**, 498-513.

Pante N., Jarmolowski A., Izaurralde E., Sauder U., Baschong W. and Mattaj I.W. (1997) *RNA*, **3**, 498-513.

Parker, R., Siliciano, P.G. and Guthrie, C. (1987) *Cell*, **49**, 229-239.

Parker, R. and Siliciano, P.G. (1993) *Nature*, **361**, 660-662.

Parkes V. and Johnston, L.H. (1992) *Nucleic Acids Res.*, **21**, 5617-5623.

Patterson, B. and Guthrie, C. (1987) *Cell*, **49**, 613-624.

Patterson, B. and Guthrie, C. (1991) *Cell*, **64**, 181-197.

Patton J. R. (1991) *Mol. Cell. Biol.* **11**, 5998-6006.

Patton, J.R. (1993) *Biochem. J*, **290**, 595-600.

Patton, J.R. (1994) *Biochemistry*, **33**, 10423-10427.

Patton, J.R., Jacobson, M.R. and Pederson, T. (1994) *Proc. Natl. Acad. Sci. USA.*, **91**, 3324-3328.

Petersen-Bjørn,S.A., Soltyk, J.D., Beggs, J.D. and Friesen, J.D. (1989) *Mol. Cell. Biol.*, **9**, 3698-3709.

Pikielny, C.W. and Rosbash, M. (1986) *Cell*, **45**, 869-877.

Pikielny, C.W., Rymond, B.C. and Rosbash, M. (1986) *Nature*, **324**, 341-345.

Plessel, G., Fisher, U. and Lührmann, R. (1994) *Mol. Cell. Biol.*, **14**, 4160-4172.

Plessel, G., Lührmann, R. and Kastner, B. (1997) *J. Mol. Biol.*, **265**, 87-94.

Plumpton, M. McGarvey, M. and Beggs, J.D. (1994) *EMBO J.*, **13**, 879-887.

Potashkin, J. and Frendewey, D. (1990) *EMBO J.*, **9**, 525-534.

Query, C.C., Bentley, R.C. and Keene, J.D. (1989) *Cell*, **57**, 89-101.

Raghunathan, P. and Guthrie, C. (1997) *RNA'97*: the second annual meeting of the RNA society.

Raghunathan, P. and Guthrie, C. (1998) *Science*, **279**, 857-879.

Raker, V.A., Plessel, G. and Lührmann, R. (1996) *EMBO J.*, **15**, 2256-2269.

Ramamurthy, L., Ingledue, T.C., Pilch, D.R., Kay, B.K. and Marluff, W.F. (1996) *Nucleic Acids Res.*, **24**, 4525-4534.

Reddy, R. and Busch, H. Small nuclear RNAs, Birnstiel ML (ed), Berlin, Springer-Verlag, 1988, pp 1-37.

Reed, R. and Maniatis, T. (1985) *Cell*, **41**, 95-105.

Reed, R. (1996) *Curr. Opin. Genet. Dev.*, **6**, 215-220.

Reyes, J.L., Kois, P., Konforti, B.B. and Konarska, M.M. (1996) *RNA*, **2**, 213-225.

Rinke, J., Appel, B., Digweed, M. and Lührmann, R. (1985) *J. Mol. Biol.*, **185**, 721-731.

Romac, J.J-M., Graff, D.H. and Keene, J.D. (1994) *Mol. Cell. Biol.*, **14**, 4662-4670.

Roscigno, R.F. and Garcia-Blanco, M.A. (1995) *RNA*, **1**, 692-706.

Roy, J., Zheng, B., Rymond, B.C. and Woolford, J.L. (1995a) *Mol. Cell. Biol.*, **15**, 445-455.

Roy, J., Kim, K.; Maddock, J.R.; Antony, J.G. and Woolford, J.L. (1995b) *RNA*, **4**, 375-390.

Ruby, S.W., Chang, T-H. and Abelson, J. (1993) *Genes. Dev.*, **7**, 1909-19

Ruby, S.W. (1997) *J. Biol. Chem.*, **272**, 17333-17341.

Rüther, U. and Muller-Hill, B. (1983) *EMBO J.*, **2**, 1791-1794.

Ruskin, B. and Green, M.R. (1985a) *Nature*, **317**, 732-734.

Ruskin, B. and Green, M.R. (1985b) *Science*, **229**, 135-229.

Ruskin, B., Zamore, P.D. and Green, M.R. (1988) *Cell*, **52**, 207-219.

Rymond, B. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 848-852.

Rymond, B., Rokeach, L.A. and Hoch S.O. (1993) *Nucleic Acids Res.*, **21**, 3501-3505.

Sachs, A., Bond, M.W. and Kornberg, R.D. (1986) *Cell*, **45**, 827-835.

Sauterer, R.A., Goyal, A. and Zieve, G.W. (1990) *J. Biol. Chem.*, **265**, 1048-1058.

Sawa, H. and Shimura, Y. (1991) *Nucleic Acids Res.*, **19**, 6819-6821.

Sawa, H. and Abelson, J. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 11269-11273.

Sawa, H. and Shimura, Y. (1992) *Genes Dev.*, **6**, 244-254.

Scherly, D., Boelens, W., Van Venrooij, W.J., Dathan, N.A., Hamm, J. and Mattaj, I.W. (1989) *EMBO J.*, **8**, 4163-4170.

Scherly, D., Boelens, W., Dathan, N.A., Van Venrooij, W.J. and Mattaj, I.W. (1990) *Nature*, **345**, 502-506.

Schmid, S.R., Linder, P. (1992) *Mol. Microbiol.*, **6**, 283-292.

Schwer, B. and Guthrie, C. (1991) *Nature*, **349**, 494-499.

Schwer, B. and Guthrie, C. (1992a) *Mol. Cell. Biol.*, **12**, 3540-3547.

Schwer, B. and Guthrie, C. (1992b) *EMBO J.*, **11**, 5033-5039.

Schwer, B. and Shuman, S. (1996) *RNA*, **2**, 574-583.

Schwer, B. and Gross, C.H. (1998) *EMBO J.*, **17**, 2086-2094.

Ségault, V., Will, C.L, Sproat, B.S and Lührmann, R. (1995) *EMBO J.*, **16**, 4010-4021.

Sellem, C.H. and Belcour, L. (1994) *Nucleic Acids Res.*, **22**, 1135-1137.

Séraphin, B., Kretzner, L. and Rosbash, M. (1988) *EMBO J.*, **7**, 2533-2538.

Séraphin, B. and Rosbash, M. (1989) *Cell*, **59**, 349-358.

Séraphin, B., Abovich, N. and Rosbash, M. (1991) *Nucleic Acids Res.*, **19**, 3857-3860.

Séraphin, B. (1995) *EMBO J.*, **14**, 2089-2098.

Siemester, G., Buchholz, C. and Hachtel, W. (1990) *Mol. Gen. Genet.*, **220**, 425-432.

Siliciano, P.G., Jones, M.H. and Guthrie, C. (1987a) *Science*, **237**, 1484-1487.

Siliciano, P.G., Brow, D.A., Roiha, H. and Guthrie, C. (1987b) *Cell*, **50**, 585-592.

Siliciano, P.G. and Guthrie, C. (1988) *Genes Dev.*, **2**, 1258-1267.

Sillekens, P.T.G., Beijer, R.C., Habets, W.J. and Van Venrooij, W.J. (1988) *Nucleic Acids Res.*, **16**, 8307-8321.

Simon, M.I.; Strathmann, M.P. and Gautam, N. (1991) *Science*, **252**, 802-808.

Singh, R., Valcarcel, J. and Green, M.R. (1995) *Science*, **268**, 1173-1176.

Shannon, K.W. and Guthrie, C. (1991) *Genes Dev.*, **5**, 773-785.

Sharp, P.A (1994) *Cell*, **77**, 805-815.

Sharp, P.A. and Burhe, C.B. (1997) *Cell*, **91**, 875-879.

Shuster, E.O. and Guthrie, C. (1988) *Cell*, **55**, 41-48.

Smith, V. and Barrell, B. (1991) *EMBO J.*, **10**, 2627-2634.

Sondek, J., Bohm, N.H., Lambright, D.G., Hamm, H.E. and Sigler, P.B. (1996) *Nature*, **379**, 369-374.

Sontheimer, E.J. and Steitz, J.A. (1993) *Science*, **262**, 1989-1996.

Sontheimer, E.J. (1994) *Mol. Biol. Reports*, **20,** 35-44.

Staknis, D. and Reed, R. (1994) *Mol. Cell. Biol.*, **14**, 2994-3005.

Steitz, J.A. (1992) *Science*, **257**, 888-889.

Strauss, E.J. and Guthrie, C. (1991) Genes Dev., 5, 629-641.

Strauss, E.J. and Guthrie, C. (1994) *Nucleic Acids Res.*, **22**, 3187-3193.

Sun, J.S. and Manley, J.L. (1995) *Genes Dev.*, **9**, 843-854.

Sun, J.S. and Manley, J.L. (1997) *RNA*, **3**, 514-526.

Takahashi, N.; Hayano, T. and Suzuki, M. (1989) *Nature*, **337**, 473-475.

Tang, J. and Rosbash, M. (1996) *RNA*, **2**, 1058-1070.

Tang, J., Abovich, N. and Rosbash, M. (1996b) *Mol. Cell. Biol.*, **16**, 2787-2795.

Tang, J., Abovich, N., Fleming, M.L., Séraphin, B and Rosbash, M. (1997) *EMBO J.*, **16**, 4082-4091.

Tani T. and Ohshima Y. (1989) *Nature*, **337**, 87-90

Tarn, W.Y., Lee, K.R. and Cheng, S.C. (1993) *Proc Natl Acad Sci U S A,* **15**; 10821-10825.

Tarn, W-Y. and Steitz, J.A. (1994a) *Genes Dev.*, **8**, 2704-2717.

Tarn, W-Y., Hsu, C.H., Huang, K.T., Chen, H.R., Kao, H.Y., Lee, K.R. and Cheng S.C. (1994b) *EMBO J.*, **13**, 2421-2431.

Tarn, W-Y., Yario, T.A. and Steitz J.A. (1995) *RNA*, **1**, 644-656.

Tarn, W-Y. and Steitz, J.A. (1996a) *Cell*, **84**, 801-811.

Tarn, W-Y. and Steitz, J.A. (1996b) *Science*, **273**, 1824-1832.

Tarn, W-Y. and Steitz, J.A. (1997) *TIBS*, **22**, 132-137.

Tazi, J., Forné, T., Jeanteur, P., Cathala, G. and Brunel, C. (1993) *Mol. Cell. Biol.*, **13**, 1641-1650.

Tazi, J., Kornstädt, U., Rossi, F., Jeanteur, P., Catala, G., Brunel, C. and Lührmann, R. (1993b) *Nature*, **363**, 283-286.

Teigelkamp, S., McGarvey, M., Plumpton, M. and Beggs, J.D. (1994) *EMBO J.*, **13**, 888-897.

Teigelkamp, S., Newman, A.J. and Beggs, J.D. (1995) *EMBO J.*, **14**, 2602-2612.

Teigelkamp, S., Whittaker, E and Beggs, J.D. (1995b) *Nucleic Acids Res.*, **23**, 320-326.

RNA 1997 Nov;3(11):1313-1326

Teigelkamp, S., Mundt, C., Achsel, T., Will, C.L. and Lührmann, R. (1997) *RNA*, **3**, 1313-1326.

Teigelkamp, S., Achsel, T., Mundt, C., Göthel, S-F., Cronshagen, U., Lane, W.S., Marahiel, M. and Lührmann, R. (1998) *RNA*, **4**, 127-141.

Terns, M.P., Lund, E. and Dahlberg, J.E. (1992) *Mol. Cell. Biol.*, **12**, 3032-3040.

Tilghman, S.M, Tiemeier, D.C., Seidman, J.G., Peterlin, B.M., Sullivan, M., Maizel, J.V. and Leder, P. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 1485-1489.

Tollervey, D. and Mattaj, I.W. (1987) *EMBO J.*, **6**, 469-476.

Triana-Alonso, F.J., Dabrowski, M., Wadzack, J. and Nierhaus, K.H. (1995) *J. Biol. Chemistry*, **11**, 6298-6307.

Umen, J.G. and Guthrie, C. (1995) *RNA*, **1**, 584-597.

Umen, J.G. and Guthrie, C. (1996) *Genetics,* **143**, 723-739.

Vaidya, V.C., Seshadri, V. and Vijayraghavan, U. (1996) *Mol. Gen. Genet.*, **250**, 267-276.

Van Dam, A. (1989) *EMBO J.*, **8**, 3853-3860.

Van Dyck, L., Jonniau, J.L., deMelo, B.T., Kleine, K. and Goffeau, A. (1994) *Yeast*, **10**, 1663-1673.

Vankan, P., McGuigan, C. and Mattaj, I.W. (1992) *EMBO J.*, **11**, 335-343.

Vijayraghavan, U., Company, M. and Abelson, J. (1986) *Genes Dev.*, **3**, 1206-1216.

Vijayraghavan, U., Parker, R., Tamm, J., Limura, Y., Rossi, J., Abelson, J. and Guthrie, C. (1986b) *EMBO J.*, **5**, 1683-1695.

Vijayraghavan, U. and Abelson, J. (1990) *Mol. Cell. Biol.*, **10**, 324-332.

Wagner, J.D.O., Company, M. and Abelson, J. (1997) *RNA'97*: the second annual meeting of the RNA society.

Wagner, J.D.O., Jankowsky, E. Company, M., Pyle, A.M. and Abelson, J. (1998) *EMBO J.*, **17**, 2926-2937.

Wall, M.A., Coleman, D.E., Lee, E., Iniguez-Lluli, J.A., Posner, B.A., Gilman, A.G. and Sprang, S.R. (1995) *Cell*, **83**, 1047-1058.

Wang, Y.; Wagner, J.D.O and Guthrie, C. (1998) *Curr. Biol.*, **8**, 441-451.

Wassarman, D.A. and Steitz, J.O. (1992) *Science*, **257**, 1918-1925.

Weidenhammer, E.M., Ruiz-Noriega, M. and Woolford, J.L.JR. (1997) *Mol. Cell. Biol.*, **17**, 3580-3588.

Wells, S.E. and Ares, M.JR. (1994) *Mol. Cell. Biol.*, **14**, 6337-6349.

Wells, S.E., Neville, M., Haynes, M., Wang, J., Igel, H. and Ares, M.J. (1996) *Genes Dev.*, **10**, 220-232.

Wersig, C. and Bindereif, A. (1992) *Mol. Cell. Biol.*, **12**, 1460-1468.

Whittaker, E., Lossky, M. and Beggs, J.D. (1990) *Proc Natl. Acad. Sci. USA*, **87**, 2216-2219.

Whittaker, E. and Beggs, J.D. (1991) *Nucleic Acids Res.*, **19**, 5483-5489.

Will, C.L., Behrens, S-E. and Lührmann, R. (1993) *Mol. Biol. Reports*, **18**, 121-126.

Will, C.L. and Lührmann, R. (1997) *Curr. Opin. Cell Biol.*, **9**, 320-328.

Wolff, T. and Bindereif, A. (1993) *Genes Dev.*, **7**, 1377-1389.

Wooley, J.C., Zukerberg, L.R. and Chung, S-Y. (1983) *Proc. Natl. Acad. Sci. USA.*, **80**, 5208-5212.

Woppmann, A., Patschinsky, T., Bringmann, P., Godt, F. and Lührmann, R. (1990) *Nucleic Acids Res.*, **18**, 4427-4438.

Woppmann, A., Will, C.L., Kornstädt, U., Zuo, P., Manley, J.L. and Lührmann, R. (1993) *Nucleic Acids Res.*, **21**, 2815-2822.

Wu, J. and Manley, J.L. (1989) *Genes Dev.*, **3**, 1553-1561.

Wu, J.Y. and Maniatis, T. (1993) *Cell*, **75**, 1061-1070.

Wu, Q. and Krainer, A.R. (1997) *RNA*, **3**, 586-601.

Wyatt, J.R., Sontheimer, E.J. and Steitz, J.A. (1992) *Genes Dev.*, **6**, 2542-2553.

Xiao, S-H. and Manley, J.L. (1997) *Genes Dev.*, **11**, 334-344.

Xie, J., Beickman, K., Otte, E. and Rymond, B.C. (1998) *EMBO J.*, **17**, 2938-2946.

Xu, Y., Petersen-Bjørn, S. and Friesen, J.D. (1990) *Mol. Cell. Biol.*, **3**, 1217-1225.

Xu, Y., Nouraini, S., Field, D., Tang, S-J. and Friesen, J.D. (1996) *Nature*, **381**, 709-713.

Yean, S-H., and Lin, R-J. (1991) *Mol. Cell. Biol.*, **11**, 5571-5577.

Yisraeli, J.K and Melton, D.A (1898) *Methods in Enz.*, **180**, 42-49.

Yu, Y-T. and Steitz, J.O. (1997) *RNA*, **3**, 807-810.

Zahler, A.M., Lane, W.S., Stolk, I.A. and Roth, M.B. (1992) *Genes Dev.*, **6**, 837-847.

Zamore, P.D. and Green, M.R. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 9243-9247.

Zamore, P.D., Patton, J.G. and Green, M.R. (1992) *Nature*, **355**, 609-614.

Zerby, D.B. and Patton, J.R. (1996) *Nucleic Acids Res.*, **24**, 3583-3589.

Zhang, M., Zamore, P.D., Carmo-Fonseca, M., Lamond, A.I. and Green, M.R. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 8769-8773.

Zhang, W-J. and Wu, J.Y. (1996) *Mol. Cell. Biol.*, **16**, 5400-5408.

Zhang, X. and Schwer, B. (1997) *Nucleic Acids Res.*, **25**, 2146-2152.

Zhou, C., Yang, Y. and Jong, A.Y. (1990) *Biotechniques*, **2**, 172-173.

Zhuang, Y. and Weiner, A.M. (1986) *Cell*, **46**, 827-835.

Zieve, G.W., Sauterer, R.A. and Feeney, R.J. (1988) *J. Mol. Biol.*, **199**, 2259-2267.

Zillmann, M,, Zapp., M.L. and Berget, S.M. (1988) *Mol. Cell. Biol.*, **8**, 814-821.

Zuo, P. and Maniatis, T. (1996) *Genes Dev.*, **10**, 1356-1368.

# ANNEX I

Sequence of the 3.5 Kb- *Xba*I fragment contained in YCpXba plasmid (Table 9) (Shannon *et al.*, 1991).

This sequence was communicated by A. Ghetti from J. Abelson's laboratory (California), it is now available in the *Saccharomyces cerevisiae* genome database of the yeast chromosome XIII; *PRP24* ORF can be found under the accession number YMR268c.

The nucleotides indicated in bold characters belong to *PRP24* ORF; the oligodeoxynucleotides used in this work are indicated by arrows.

```
1   TCTAGATGTG GATGAATTGA ACATTCCAAA CGAAATATCC CGGAATATCT

51  TAGTTAAACT GGACAGCTTA TTCGAAGGGT TGCATGATAA AATAGCCAAA

101 GAGAACGAAT TCGATGTAAG GCAAGATAAG CATTCTAATA ATATAAGGGC

151 GAATCAGATA GACGACGAGC CCATGCAGGC CAATCGACGA ATTAAATACA

201 CATATCATGA CTTGGTTTCA AGAGGTTGTG AAATGAACGA GGATATGACG

251 GATATCTACA TGAAATCATT AGAGCTATAT AACGATATCC CAGAAAAGTA

301 TAAAAAGAGA AAATTCAGAC TGCCCAAACA AATTCTAAAA AAGTATCACC

351 AGCCAAAAAA GACCAGTTCA TACTTAAAGG AGCTATTAAG TAAAACAAGA

401 GAAGATTTCA TACCTGTGGA AAAGTTATTG AAGGATAAAC GGTTGACATC

451 AAAGGACAAA TCCAAGCTAC AACGCTTAAA TCGTGAAGAG ACGGAAGATG

501 CTCTAAACAA AAGAACATTT TTTCAGGTAA AGGGCTACTT AGAGGATGAG

551 AATGAGATCT CTGATTACGA GTTGGACGAC TGCCTCATAG AACTGCCTAA

601 TGGGAACATA TGAATGAGAA AATCTTTTTT TTACCTTTTT CCATTAAGAA

651 ATATTCATCT CCTTTTAAGA TCAACATATG TAACAATAAA TATACTCTTT

701 TAAAGATGTT CTTTTAATTT ACTTACTGAA ATTTAATATT AATAAAAGTT

751 CGAAAACAAA GACTAAATAC AAAAATAAAA AGATCATAAA TTCAAACTTT

801 TAAGAATGTC TTTTACGGAT ACATATTTAC ATATTAGACT ATGATAATCA

851 TATATGCTCC TTTCTATCTC TTCTTGCTGA ACTTTCTTCC TTTTTATGTT
```

901 TTTTTGACTT TTTGGATTTT TTATGCTTCG ATTTATCACC TTCATCGCTG

951 CTATGCTTAC TTTTCTTTCC TTTTTTACTC TCCTTCTTGC TTTTCTTCTT

1001 CTTCTTCTTC TTACTCTCAC TGTCGTTACT AGTCTTTTTA TCTTTTTTCA

1051 ATTTTTTTCT CTTGACCTTG TTATCATCTT CATCACGCCT TCTTCTCTTC

1101 TTCCCTTTAC TCGAACTTGC ACTAGATACA ACAAAGCTGG CTTCCTCCTT

1151 TTTACCAAGA TTAGTAATGG TTCCTTTCAG CCCTTCTCCC TTTACAAACC

1201 ACCTGTATAG AGGTGAGCTA CTTTTCGATA CAGCAGTAGC AACTGCCTCA

1251 TTTTGAGTAA ATTTAATACT GCCATTATTC GAATCAGTGC TTACATCCAG

1301 GTTCTTCAGA TGTCCATCAA ATAGCCTTTC CCACCATGCT TCGCCATCAT

1351 TCCCGCCAGG AGCATTACCC AATCCCTTCT TATCTCTCTT GTGCTTTACC

1401 AGTATGGGTC TTTTCAAACC ACCTTCTCTA AACGCTTCTC CTTCTTTCCA

1451 ACCATATGAT ATCAGATATT CTTTGCTATC CATGTCCGAT TTATTCTGAT

1501 AATATTTCCG CTTGTATCAT TCAATTTTCT ATGTACATCA TGACAATGAA

1551 TTGTAGAGTG CATCAATGTC TTTTTTAAAG TCATCGCTAA AAAATTTTCA

1601 GGCGGGAACT TATTCAAAAT GGCAGATTTA CTACAACATT GAAAACCCTT

1651 TATATTTATT ATATGAGAAT CATTAGCGAT AAACAAGCCC ACATATTCAC

1701 TTCATAATCT GTGATA**ATGG AGTATGGACA TCACGCTAGA CCAGATTCAA**

S8948

1751 **AACGACCATT GGATGAAGGA TCACCAGCAG CCGCAGGATT AACTTCTAAG**

1801 **AAGGCGAACG AAGCTTTAAC AAGAAATAGA GAATTAACAA CTGTATTAGT**

P5118

1851 **CAAAAATTTG CCTAAGAGCT ATAATCAAAA CAAAGTCTAT AAGTACTTCA**

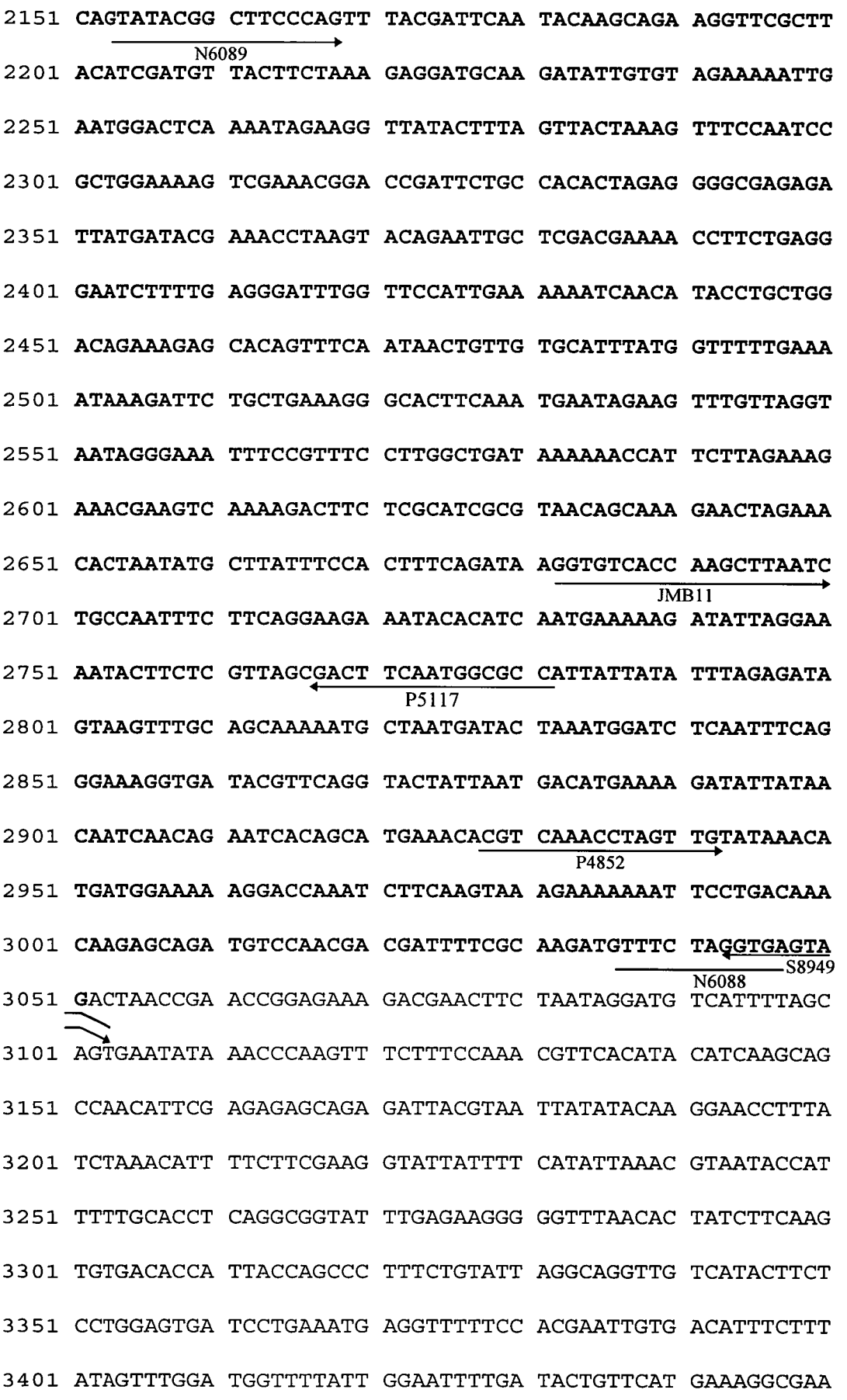
1901 **AACATTGTGG GCCCATTATA CACGTTGATG TTGCGGATTC GCTGAAGAAG**

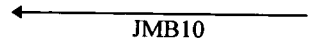
1951 **AACTTTCGTT TTGCACGTAT TGAATTTGCC AGGTATGATG GAGCCCTCGC**

2001 **TGCAATAACC AAAACACACA AAGTTGTAGG TCAGAATGAA ATTATAGTAT**

2051 **CTCATTTAAC AGAATGCACA TTATGGATGA CGAATTTCCC CCCAAGTTAT**

2101 **ACTCAAAGAA ATATTAGAGA TCTATTGCAA GATATTAACG TTGTTGCACT**


2151 CAGTATACGG CTTCCCAGTT TACGATTCAA TACAAGCAGA AGGTTCGCTT
N6089
2201 ACATCGATGT TACTTCTAAA GAGGATGCAA GATATTGTGT AGAAAAATTG
2251 AATGGACTCA AAATAGAAGG TTATACTTTA GTTACTAAAG TTTCCAATCC
2301 GCTGGAAAAG TCGAAACGGA CCGATTCTGC CACACTAGAG GGGCGAGAGA
2351 TTATGATACG AAACCTAAGT ACAGAATTGC TCGACGAAAA CCTTCTGAGG
2401 GAATCTTTTG AGGGATTTGG TTCCATTGAA AAAATCAACA TACCTGCTGG
2451 ACAGAAAGAG CACAGTTTCA ATAACTGTTG TGCATTTATG GTTTTTGAAA
2501 ATAAAGATTC TGCTGAAAGG GCACTTCAAA TGAATAGAAG TTTGTTAGGT
2551 AATAGGGAAA TTTCCGTTTC CTTGGCTGAT AAAAAACCAT TCTTAGAAAG
2601 AAACGAAGTC AAAAGACTTC TCGCATCGCG TAACAGCAAA GAACTAGAAA
2651 CACTAATATG CTTATTTCCA CTTTCAGATA AGGTGTCACC AAGCTTAATC
JMB11
2701 TGCCAATTTC TTCAGGAAGA AATACACATC AATGAAAAAG ATATTAGGAA
2751 AATACTTCTC GTTAGCGACT TCAATGGCGC CATTATTATA TTTAGAGATA
P5117
2801 GTAAGTTTGC AGCAAAAATG CTAATGATAC TAAATGGATC TCAATTTCAG
2851 GGAAAGGTGA TACGTTCAGG TACTATTAAT GACATGAAAA GATATTATAA
2901 CAATCAACAG AATCACAGCA TGAAACACGT CAAACCTAGT TGTATAAACA
P4852
2951 TGATGGAAAA AGGACCAAAT CTTCAAGTAA AGAAAAAAAT TCCTGACAAA
3001 CAAGAGCAGA TGTCCAACGA CGATTTTCGC AAGATGTTTC TAGGTGAGTA
S8949
N6088
3051 GACTAACCGA ACCGGAGAAA GACGAACTTC TAATAGGATG TCATTTTAGC
3101 AGTGAATATA AACCCAAGTT TCTTTCCAAA CGTTCACATA CATCAAGCAG
3151 CCAACATTCG AGAGAGCAGA GATTACGTAA TTATATACAA GGAACCTTTA
3201 TCTAAACATT TTCTTCGAAG GTATTATTTT CATATTAAAC GTAATACCAT
3251 TTTTGCACCT CAGGCGGTAT TTGAGAAGGG GGTTTAACAC TATCTTCAAG
3301 TGTGACACCA TTACCAGCCC TTTCTGTATT AGGCAGGTTG TCATACTTCT
3351 CCTGGAGTGA TCCTGAAATG AGGTTTTTCC ACGAATTGTG ACATTTCTTT
3401 ATAGTTTGGA TGGTTTTATT GGAATTTTGA TACTGTTCAT GAAAGGCGAA




3451 ACTATTTAAT GGCTTCCCTG CTGGAACTTT ATACTTTCTA AACCACTCCC

3501 TAGTCGTA