IDENTIFICATION AND CHARACTERISATION OF PRP45P AND PRP46P, TWO NOVEL PRE-mRNA SPLICING FACTORS IN *SACCHAROMYCES CEREVISIAE*
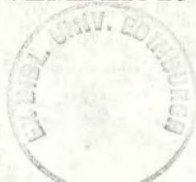
MICHAEL ALBERS

A THESIS PRESENTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF EDINBURGH

NOVEMBER 1999

ACKNOWLEDGEMENTS

Nuclear pre-mRNA splicing, the removal of intervening sequences from pre-messenger RNA, proceeds via two distinct, sequential transesterification reactions. A large and highly dynamic RNA-protein complex, termed the spliceosome, is required to catalyse these reactions. Despite years of extensive research, the identification of the spliceosomal components is far from complete. Although all participating RNA molecules are known, the number of identified protein factors is still continuously increasing.

We set out to investigate protein-protein interactions within the spliceosome of the budding yeast *Saccharomyces cerevisiae* by performing exhaustive two-hybrid screens using well characterised splicing proteins as baits. This approach should allow the identification of i) novel splicing proteins, ii) novel interactions between known splicing factors and iii) links between splicing and other cellular pathways, especially processes of mRNA metabolism.

Prp22p is an RNA helicase with at least two distinct functions in the splicing pathway. It is required for the second transesterification reaction to proceed and in addition it has a role during spliceosome disassembly. When Prp22p was used as bait in a two-hybrid screen, the most statistically significant interacting protein found was the Fun20 protein.

The 42 kDa Fun20 protein was previously shown to be essential for cell viability, but no further characterisation had been performed (Fun = **f**unction **un**known). In order to investigate whether the protein plays a role in pre-mRNA splicing, a strain was generated carrying a protein A-tagged and conditionally regulated *FUN20* gene. Growing the strain under non-permissive conditions leads to an accumulation of pre-mRNA within the cell, showing that the Fun20 protein is indeed required for splicing *in vivo*. The protein was renamed Prp45p to indicate its role in **pr**e-mRNA **p**rocessing. Using the epitope-tagged version of the protein in coimmunoprecipitation experiments, it was found that Prp45p coprecipitates pre-mRNA, splicing reaction-intermediates, the spliced exons and the excised intron from cell extracts. This strongly suggests that Prp45p is associated with the spliceosome

throughout the splicing reactions. Furthermore the tagged-protein weakly coprecipitates the U2, U5 and U6 snRNAs from cell extracts, showing its association with a subset of spliceosomal snRNPs. Depletion of Prp45p from cell extracts completely abolishes splicing of added actin pre-mRNA in these extracts. However, splicing activity can be at least partially restored by adding back recombinant His-tagged Prp45p, that had been produced in *Escherichia coli* and affinity purified. This shows that Prp45p is required for splicing *in vitro* and furthermore suggests a role for the protein before the first transesterification takes place.

Prp45p was then used as bait in a two-hybrid screen and interacted significantly with a protein of unknown function encoded by ORF *YPL151c*. Based on its amino acid sequence, the 50 kDa Ypl151 protein can be grouped into a subfamily of WD proteins with nuclear localisation, which share extensive homology to the pleiotropic regulator proteins PRL1 and PRL2 in *Arabidopsis thaliana*. *YPL151c* is essential for cell viability, as shown by deletion of the ORF from the genome. A strain with a regulated version of the *YPL151c* gene was constructed and it was demonstrated that depletion of the protein from the cells leads to a severe splicing deficiency. Ypl151p has been renamed Prp46p, to indicate its function in pre-mRNA splicing.

Therefore, through the use of exhaustive two-hybrid screens, two novel pre-mRNA splicing factors have been identified, that probably interact with each other. Some functional characterisation has been performed to determine at which stage of the splicing process they might act.

Common abbreviations

AD	Activation Domain
Amp	Ampicillin
ARS	Autonomous Replication Sequence
ATP	Adenosine 5'-triphosphate
BD	(DNA) Binding Domain
β-Gal	β-Galactosidase
BLAST	Basic Local Alignment Research Tool
bp	Base pair
BSA	Bovine Serum Albumin
°C	Degrees Celcius
CC	Commitment Complex
CEN	Centromere Sequence
cDNA	Complementary Deoxyribonucleic acid
cm	Centimetre
CTP	Cytidine 5'-triphosphate
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
ddNTP	Dideoxynucleoside triphosphate
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediaminetetraacetic acid
5-FOA	5-Fluoro-orotic acid
g	Gram(s)
Gal	Galactose
GTP	Guanidine 5'-triphosphate
h	Hour(s)
HEPES	N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid
HRP	Horse Radish Peroxidase
kb	Kilobase
kDa	KiloDalton
L	Litre(s)
LI	Lariat-intron
LI-E2	Lariat-intron exon2
M	Molar
mA	Milliampere(s)
mg	Milligram(s)
mM	Millimolar
mm	Millimetre(s)
min	Minute(s)
ml	Millilitre(s)
mRNA	Messenger Ribonucleic Acid
MW	Molecular Weight

NCBI	National Center for Biotechnology Information
ng	Nanogram(s)
nm	Nanometre(s)
nt	Nucleotide
OD_{260}	Optical Density at 260 nm
OD_{280}	Optical Density at 280 nm
OD_{600}	Optical Density at 600 nm
OLB	Oligonucleotide Labelling Buffer
ORF	Open Reading Frame
p	Plasmid designation
PAGE	Polyacrylamide Gel Electrophoresis
PAS	Protein A-Sepharose
PCR	Polymerase Chain Reaction
PEG	Polyethylene Glycol
pmol	Picomole(s)
rDNA	Ribosomal DNA
RNA	Ribonucleic Acid
RNase	Ribonuclease
rpm	Revolutions per minute
SDS	Sodium Dodecyl Sulphate
SET	Sucrose/EDTA/Tris
SGD	*Saccharomyces* Genome Database
snoRNA	Small Nucleolar Ribonucleic Acid
snRNA	Small Nuclear Ribonucleic Acid
snRNP	Small Nuclear Ribonucleoprotein Particle
SSC	Sodium Chloride/Sodium Citrate Buffer
TAE	Tris/Acetate (buffer)
TBE	Tris/Borate/EDTA (buffer)
TBS-TT	Tris-buffered Saline/Tween 20/Triton-X-100
TE	Tris/EDTA (buffer)
TEMED	*N,N,N',N'*-tetramethyl-ethylenediamine
Tet	Tetracycline
tRNA	Transfer Ribonucleic Acid
Tris	Tris(hydroxymethyl)aminomethane
µg	Microgram(s)
µl	Microlitre(s)
U	Unit(s)
UAS	Upstream Activating Sequence
UTP	Uridine triphosphate
UTR	Untranslated region
UV	Ultra Violet
V	Volt(s)
v/v	Volume per unit volume
w/o	Without
w/v	Weight per unit volume
X-Gal	5-Bromo-4-Chloro-3-Indolyl-β-*D*-Galactoside
YMM	Yeast Minimal Media
YPDA	Yeast/Peptone/Dextrose/Adenine (media)

Amino acid abbreviations

Amino Acid	3-letter code	1-letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartate	Asp	D
Cysteine	Cys	C
Glutamate	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter I - Introduction

Chapter II - Materials and Methods

Chapter III - The Slu7p and Prp18p two-hybrid screens

Chapter IV - Characterisation of prey proteins found in the Slu7p and Prp18p two-hybrid screens

Chapter V - The Prp22p two-hybrid screen

Chapter VI - Prp45p is a novel pre-mRNA splicing factor

Chapter VII - Prp46p is required for pre-mRNA splicing in vivo

Chapter VIII - Overview and future work

References

Chapter I

Introduction

I.1 Nuclear pre-mRNA splicing

In eukaryotes the primary transcripts of most nuclear protein-coding genes are interrupted by intervening sequences (introns) which are removed from the pre-mRNA in a process designated as pre-mRNA splicing. The protein coding regions (exons) become joined to yield the mature message which is subsequently exported into the cytoplasm where its information is read and translated into protein. Pre-mRNA splicing is an essential process for gene expression since the intron sequences in most cases contain in-frame stop codons which would cause premature termination of translation, leading to the production of shortened, non-functional proteins. For genes containing multiple introns, pre-mRNA splicing can also provide a means for regulation of gene expression. By alternative splicing a gene can be switched on and off or different isoforms of a protein can be generated from the same pre-mRNA in e.g. a cell cycle or tissue specific manner (for a review on alternative splicing see Valcarcel *et al.*, 1995). In higher eukaryotes, a gene can contain more than 50 introns and the introns can be of immense size (up to 200,000 nucleotides in length).

In the unicellular eukaryote *Saccharomyces cerevisiae*, which will be referred to as yeast in this work, gene organisation is considerably simpler. Only relatively few, approximately 4% of the 6000 or so yeast genes, have introns and those which have, contain only one close to the 5' end of the transcript. Exceptions are the *MATa*1 and the ribosomal *RPL8A* genes, which both have two small introns. Nevertheless, splicing is an essential process in yeast, since many of the genes that contain introns are essential. Furthermore, a lot of highly transcribed genes (among which are many ribosomal protein genes) are interrupted. In fact, normally about 30% of all transcript present in a yeast cell is derived from intron-containing genes, illustrating the importance of the splicing process.

The investigation of the mechanism of pre-mRNA splicing was facilitated greatly by the development of *in vitro* assays for both the mammalian and the yeast system (Krainer *et al.*, 1984; Padgett *et al.*, 1984; Ruskin *et al.*, 1984; Lin *et al.*, 1985). In addition, the application of numerous genetic approaches in yeast have contributed a great deal to the understanding of the splicing process (reviewed in Rymond and Roshbash, 1992; Beggs, 1995). It was revealed that introns are removed from nuclear

pre-mRNA in a two-step splicing reaction that is conserved in eukaryotes from yeast to man.

I.1.1 Pre-messenger RNA structure

Surprisingly, there are only three short conserved sequence elements within the introns, which are important but not sufficient for accurate and efficient splicing. These define the 5' splice site, the 3' splice site and the so-called branchpoint sequence. In mammalian pre-mRNAs these sequence elements are less conserved than in yeast (figure I.1), probably reflecting differences in the associated splicing machinery. A fourth element, which is less pronounced in yeast is a pyrimidine-rich region upstream of the 3' splice site.
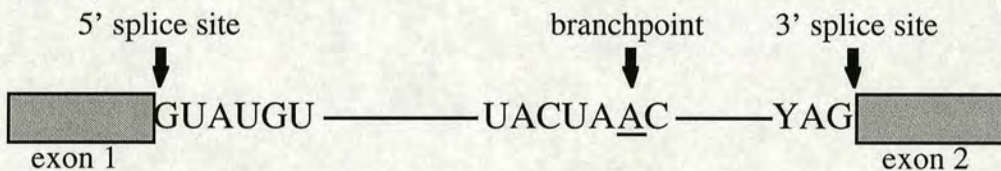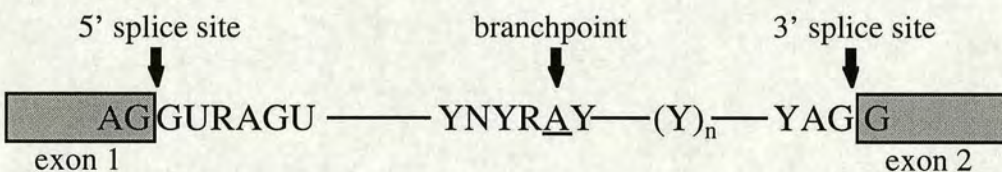
Figure I.1:
Schematic representation of a yeast and a mammalian intron. The three conserved sequence elements, the 5' splice site, the branchpoint sequence and the 3' splice site are indicated. The branchpoint adenosine is underlined. Exons are depicted as grey boxes, the intron as a line. $(Y)_n$ indicates the polypyrimidine tract. Y - pyrimidine, R - purine, N - any base.

The importance of these elements has been shown through site-directed mutagenesis and through the analysis of naturally occurring mutations within these regions both *in vivo* and *in vitro* (for a review see Nilsen, 1998 and references therein). It has to be noted, however, that additional cis-elements such as the exon sequences flanking the intron, sequences downstream of the 5' splice site as well as the intron length and the distance between the branchpoint sequence and the 3' splice site influence splicing efficiency (Pikielny and Roshbash, 1985; Newman and Norman, 1991, 1992; Luukkonen and Seraphin, 1997; Puig *et al.*, 1999).

The low information content within pre-mRNA introns is probably compensated by the interaction of numerous trans-acting factors, both RNAs and proteins, with the substrate to achieve accurate and efficient splicing.

I.1.2 The splicing reaction - a two step mechanism

Introns within nuclear pre-mRNAs are removed by two successive transesterification reactions (figure I.2). The first step entails cleavage of a phosphodiesterbond at the 5' splice site to yield a free 5' exon intermediate (exon 1) with a free 3'-hydroxyl group. Concomitantly, the 5' end of the intron is joined to a 100% conserved adenosine residue within the branchpoint sequence via an unusual 2'-5' phosphodiester bond. This forms the second intermediate of the splicing reaction, a branched intron attached to the 3' exon (often referred to as lariat-intron exon 2 structure). In the second step, the lariat-intron exon 2 structure is cleaved at the 3' splice site and the two exons become ligated via a 3'-5' phosphodiesterbond. This gives the products of the splicing reaction: the spliced exons and the excised lariat-intron with a 3'-hydroxyl terminus. By incorporating modified nucleotides at the splice site junctions it was demonstrated that chemically, both transesterifications are S_N2 reactions (Maschhoff and Padgett, 1993; Moore and Sharp, 1993). Thus, splicing of nuclear pre-mRNA resembles mechanistically the removal of group II introns from the primary transcript in the genomes of organelles in lower eukaryotes, plants and also bacteria (reviewed in Michel and Ferat, 1995). Since group II introns have been shown to splice autocatalytically *in vitro*, i.e. the catalytic activity lies within the intron itself,

4
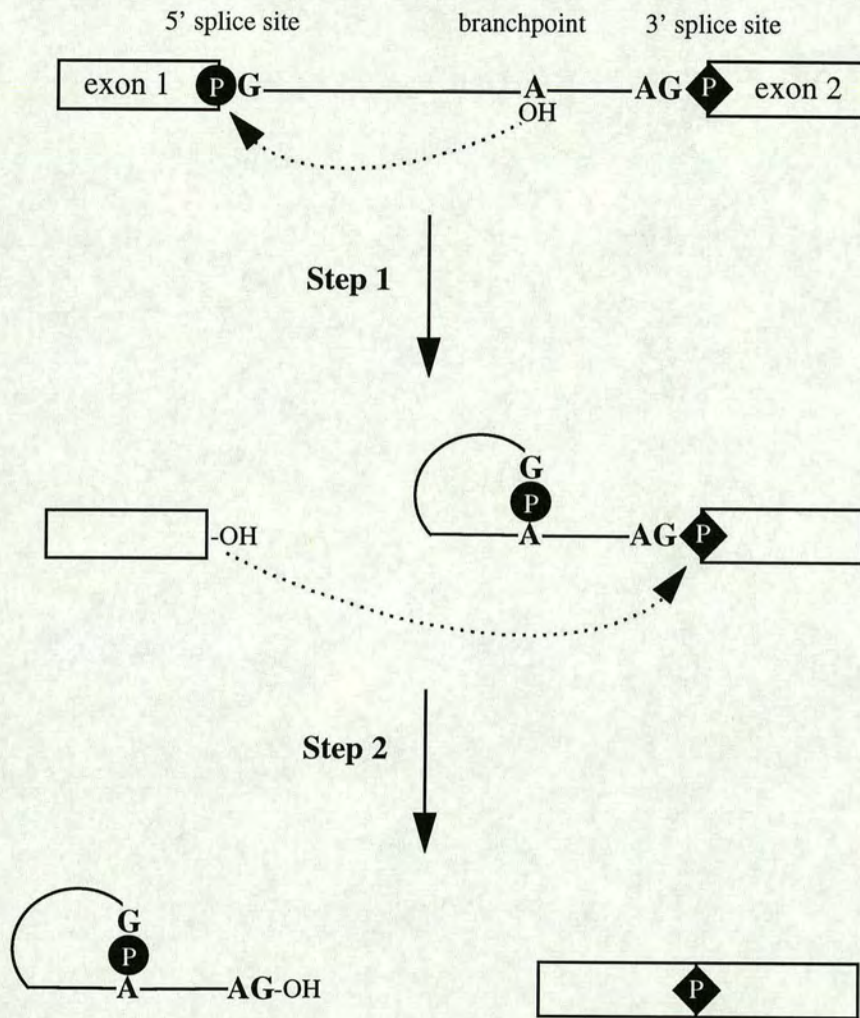
Figure I. 2:
Schematic representation of the pre-mRNA splicing reactions. Exon sequences are represented as boxes, the intron as a line. The dotted arrows indicate the sequential nucleophilic attacks of free hydroxyl groups on the phosphate groups (P) at the 5'- and 3' splice sites, respectively.

an RNA-based catalysis has been proposed also for nuclear pre-mRNA splicing. However, nuclear pre-mRNA introns obviously do not fulfil the sequence requirements needed to allow autocatalytic splicing and thus numerous trans-acting factors are required to promote or support structural rearrangements (some of which consume ATP) to juxtapose the reactive nucleotides within the catalytic centre of the spliceosome (for review see e.g. Jacquier, 1990; Guthrie, 1991; Newman, 1994).

I.2 The spliceosome - a dynamic ribonucleoprotein complex

A myriad of trans-acting proteins and RNAs are required to excise introns from nuclear pre-mRNA. The majority of these factors assemble into five small nuclear ribonucleoprotein particles (snRNPs), which again assemble onto the substrate RNA in a highly ordered fashion to form an even larger complex, termed the spliceosome. The snRNPs, named U1, U2, U4, U5 and U6, each contain an essential snRNA molecule, the U1, U2, U4, U5 and U6 snRNAs, respectively, and, with the exception of the U6 snRNP, share a number of small common core proteins. The U6 snRNP differs considerably with respect to biogenesis and basic composition from the other snRNPs (see below). In addition, a number of snRNP specific proteins have been identified, which are specifically associated with only one particular snRNP. Finally, a still increasing number of non-snRNP proteins contributes to the formation of the complex, its catalytic activity and its disassembly. A very recent, comprehensive overview on the spliceosomal components in yeast as well as in mammals, including a list of the protein factors involved, is given by Burge *et al.* (1999).

I.2.1 Biogenesis and general composition of the snRNP particles

After transcription by RNA polymerase II and export into the cytoplasm, the human U1, U2, U4 and U5 snRNAs become tightly bound by a heteromeric complex of seven small core proteins B, D1, D2, D3, E, F and G. Binding of the core proteins to the snRNAs plays a major role in the biogenesis and the transport of the snRNPs

into the nucleus. The seven core proteins are collectively referred to as Sm proteins and were originally identified as targets of autoantibodies from patients suffering from systemic lupus erythematosus (Lerner and Steitz, 1979). All Sm proteins share a short conserved region of aminoacid residues, which was thus designated as the Sm motif (Cooper *et al.,* 1995; Hermann *et al.,* 1995; Seraphin, 1995). The Sm proteins bind to a short single-stranded sequence element, the Sm site, present in the snRNAs (Branlant *et al.,* 1982; Guthrie and Patterson, 1988). The Sm site is sufficient for binding of the Sm proteins to the snRNA, however, it was demonstrated, that the efficiency of binding is influenced by snRNP specific proteins and cis-elements in the snRNA (Jarmolowski and Mattaj, 1993; Nelissen *et al.,* 1994). Importantly, the Sm proteins promote the hypermethylation of the pre-mRNA cap structure and constitute, together with the cap, a nuclear localisation signal that is essential for transport of the snRNPs into the nucleus, where their maturation continues with the addition of snRNP specific proteins (Zieve and Sauterer, 1990; Plessel *et al.,* 1994).

This pathway of snRNP biogenesis is highly conserved in yeast, since a homologous set of Sm proteins has been identified and for some of the proteins functional homologies have been demonstrated (Rymond *et al.,* 1993; Bordonne and Tarassov, 1996; Neubauer *et al.,* 1997; Gottschalk *et al.,* 1998).

The distinct biogenesis of the U6 snRNP starts with the generation of its snRNA: the U6 snRNA is transcribed by RNA polymerase III, differs in its cap structure from the other spliceosomal snRNAs and furthermore does not contain an Sm site. The U6 snRNA is therefore not recognised by the canonical Sm core proteins and also does not become hypermethylated. As a consequence, it is thought, and at least for higher eukaryotes it has been demonstrated, that the U6 snRNA is retained largely in the nucleus (reviewed in Reddy and Busch, 1988). However, recently proteins have been identified which also contain Sm motifs and which (apart from one) associate with the U6 snRNA, comparably to the association of the canonical Sm proteins with U1, U2, U4 and U5 snRNAs. Indeed, based on sequence similarity, coimmunoprecipitation experiments and extensive two-hybrid analyses, it has been proposed that these proteins form a complex analogous to the complex of Sm proteins, but which associates with U6 snRNA. Hence these proteins have been named Lsm

(for like Sm) (Fromont-Racine *et al.,* 1997; Mayes *et al.,* 1999; Salgado-Garrido *et al.,* 1999). The role of the Lsm proteins for the biogenesis of the U6 snRNP is not yet entirely clear, although is has been demonstrated that most of them stabilise the U6 snRNA, maybe by promoting U4/U6 snRNA association and later on also the formation of the U4/U6-U5 tri-snRNP particle (see below; Mayes *et al.,* 1999). However, not all of the Lsm proteins are essential for splicing (such as Lsm1p), although they are found complexed with the other Lsm proteins, which has led to the suggestion that the Lsm proteins could form different complexes with distinct functions. For instance, a role in mRNA decapping has been proposed for Lsm1p (and associated proteins) (Boeck *et al.,* 1998).

The biogenesis of the snRNPs is then completed by the addition of snRNP specific proteins onto the snRNP/core protein complex. In many cases it has been demonstrated that the snRNP specific proteins are transported into the nucleus independently from the snRNA/core protein complex, suggesting that the addition of the snRNP specific proteins occurs predominantly in the nucleus (e.g. Kambach and Mattaj, 1992; Romac *et al.,* 1994). In yeast, many of the snRNP specific proteins have been identified by genetic means, and their association with a particular snRNP has been demonstrated afterwards by coprecipitation of snRNAs together with the proteins (see Beggs, 1995 and references therein). In recent years, affinity purification of snRNPs using well characterised and tagged snRNP specific proteins and subsequent identification of the associated proteins by mass spectrometry, accelerated the identification of the snRNP components (Neubauer *et al.,* 1997; Gottschalk *et al.,* 1998; Caspary *et al.,* 1999; Stevens and Abelson, 1999).

I.2.2 The spliceosome assembly cycle

In the nucleus, the snRNPs assemble in a highly ordered manner onto the pre-mRNA to form a spliceosome capable of performing the two transesterification reactions. Major conformational rearrangements take place within the snRNPs as well as between snRNPs during the formation of an active spliceosome. These structural

changes also require the action of numerous non-snRNP proteins, some of which will be discussed in detail later (section I.4). A very simplified view of the order of events during the formation of an active spliceosome and its disassembly is depicted in figure I.3.

The earliest specific event in the assembly of the spliceosome is the formation of the so-called commitment complex, i.e. binding of the U1 snRNP to the 5' splice site (Legrain *et al.*, 1988; Seraphin and Rosbash, 1989). A short, conserved sequence at the 5' end of the U1 snRNA base pairs with the 5' splice site sequence in an ATP-independent manner (Seraphin and Rosbash, 1989). In yeast the interaction of the U1 snRNP with the pre-mRNA promotes subsequent binding of the U2 snRNP to the branch site. The U2 snRNP addition is ATP-dependent as are all subsequent steps of the splicing pathway. Nucleotides within the U2 snRNA base pair with the branch site sequence UACUAAC, but leave the branchpoint adenosine bulged out of the short helix. Therefore, the adenosine is available as the nucleophile for the attack of the phosphate at the 5' splice site later, to initiate the first transesterifiacton reaction (Parker *et al.*, 1987; Query *et al.*, 1994). The subsequent addition of a pre-formed tri-snRNP particle, assembled through association of a U4/U6 di-snRNP with the U5 snRNP, leads to the formation of the active spliceosome. Many ATP-dependent rearrangements within this complex take place before and during the course of the transesterification reactions that involve the action of numerous non-snRNP proteins (reviewed in Kraemer, 1995; Nilsen, 1998 and references therein). After completion of the splicing reactions, the spliced messenger RNA is released, the spliceosome disassembles and the snRNPs and non-snRNP proteins can assemble onto another pre-mRNA molecule. The intron-lariat becomes debranched and degraded.

Figure I.3:
Schematic representation of the spliceosome assembly pathway. U1 snRNP binds to the 5' splice site (5'ss) to form the commitment complex. U2 binds to the branchpoint sequence (BP) to form the pre-spliceosome. A pre-formed tri-snRNP particle consisting of the U4/U6-U5 snRNPs associates to form the active spliceosome, which then performs the splicing reaction. After completion of the reaction, the mRNA is released from the spliceosome, the components which disassemble and are recycled. The intron-lariat product is debranched and the RNA degraded.
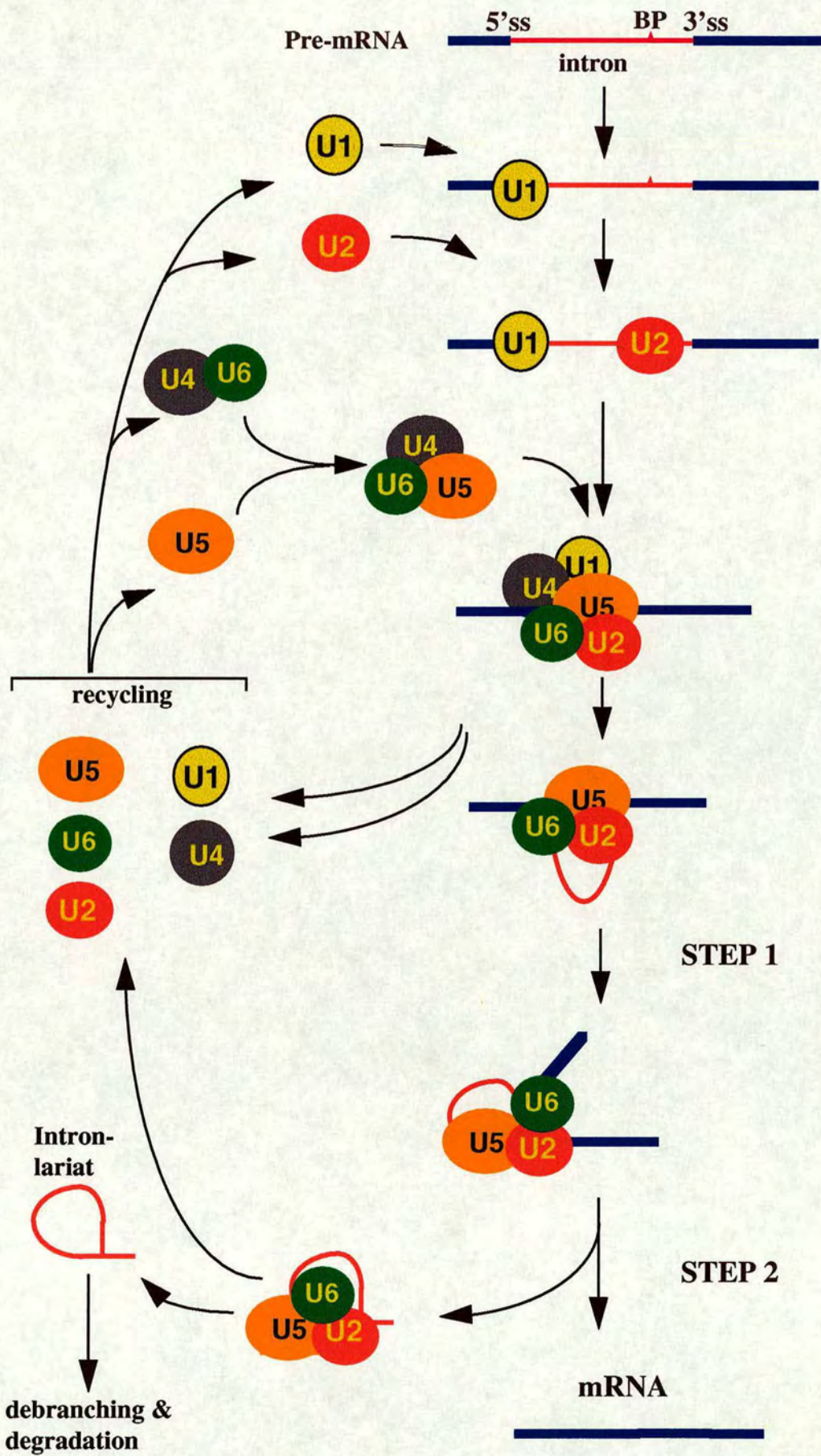
I.3 Protein actions during spliceosome assembly and splicing

This section will describe in more detail the complex actions of snRNP- and non-snRNP proteins toward the assembly of an active spliceosome. The intention here is not to give a complete listing of the protein factors involved, but to give an idea about the complexity of protein-protein, protein-RNA as well as RNA-RNA interactions. The focus will mainly lie on factors that interact with important elements in the substrate RNA and on proteins that contribute to contacts between the snRNPs. The yeast system will form the basis of this summary, however, where appropriate, comparisons with the mammalian system will be made.

I.3.1 Formation of the commitment complex

As mentioned above, it was known for some time that base pairing of U1 snRNA to the 5' splice site plays an important role for the formation of the commitment complex (e. g. Seraphin and Rosbash, 1989). But until recently no direct protein contacts had been defined within this region of the pre-mRNA during this early stage of spliceosome formation. By using pre-mRNAs with 4-thiouridine-substituted 5' splice sites (and surrounding regions) in crosslinking experiments, Zhang and Rosbash (1999) were able to demonstrate that at least 7 of the 16 proteins tightly associated with the U1 snRNP bind to the pre-mRNA substrate. Among those are three core proteins SmB, SmD1, SmD3 and the U1 snRNP specific proteins Snp1p, Yhc1p, Snu56p and Nam8p (Raker *et al.*, 1996; Smith and Barrel, 1991; Tang *et al.*, 1997; Gottschalk *et al.*, 1998). The human homologues of both Snp1p and Yhc1p, U1-70K and U1-C, have been shown to directly bind to loop 1 of the U1 snRNA (Hamm *et al.*, 1988). For Snp1p this has also been demonstrated in yeast (Kao and Siliciano, 1992). Depletion of Yhc1p from the yeast cells does not effect U1 snRNP levels or U1 snRNP assembly, but in the absence of functional Yhc1p, commitment complex (CC) formation *in vitro* is seriously inhibited (Tang *et al.*, 1997). These data suggested, that Snp1p and Yhc1p might help in the formation or stabilisation of the short U1 snRNA/5' splice site helix. Indeed crosslinking of Yhc1p was observed

exclusively within the 5' splice site sequence within the intron but not with exon sequences. Snp1p, however, crosslinked with exon sequences upstream of the 5' splice site, but not with intron sequences, suggesting that its function during CC formation does not include a direct binding of the U1 snRNA/5' splice site helix. The crosslinking studies by Zhang and Rosbash (1999) identified Snu56p and Nam8p as the first U1 snRNP associated proteins to bind to non-conserved regions downstream of the 5' splice site. Additional characterisation of Nam8p showed that it is non-essential for cell viability and that it is not required for CC formation in a wild-type context. However, CC formation on uncapped pre-mRNA or on pre-mRNA with a mutant 5' splice site was significantly inhibited in cell extracts depleted of Nam8p (Puig et al., 1999), suggesting that Nam8p is required for the stabilisation of the commitment complex. (It is known that the cap binding complex (CBC) directly contacts the U1 snRNP and thereby stabilises the U1/pre-mRNA interaction (e.g. Fortes et al., 1999)). Puig et al. (1999) demonstrated furthermore by introducing substitutions into the Nam8p binding site of the pre-mRNA that alternative 5' splice site choice and splicing efficiency is effected in a Nam8p dependent manner. Therefore it seems that early 5' splice site recognition is established not only by U1 snRNP addition onto the 5' splice site of the pre-mRNA, but that other regions of the intron and additional protein factors that bind to these non-conserved regions have some influence.

It was noted 10 years ago that in acrylamide gels, two commitment complexes (CC1 and CC2) can be resolved (Seraphin and Rosbash, 1989). For formation of CC1 the 5' end of the U1 snRNA as well as an intact 5' splice site was shown to be required. CC2, however, only forms when in addition to an intact 5' splice site, an intact branch site sequence is also present (Seraphin and Rosbash, 1991). These data suggested that apart from the U1 snRNP and associated proteins, another factor (or other factors) binds to the branchpoint prior to U2 snRNP addition. A good candidate for this association was identified as the Mud2 protein (Abovich et al., 1994). Originally isolated as a mutant synthetic lethal with a deletion in the U1 snRNA, it was demonstrated in immunoprecipitation experiments that Mud2 is not stably associated with the U1 snRNP, but that it could be coprecipitated from CC2 in the presence of pre-mRNA. By introducing a mutation into the branch site sequence, the Mud2 protein

could no longer be co-precipitated, suggesting that it indeed binds to the branch site. When the non-essential Mud2 protein was used in a synthetic lethal screen, another protein was identified which was subsequently shown to bind to the branch site sequence during CC2 formation: the branchpoint bridging protein (BBP) (Abovich and Rosbash, 1997; Berglund *et al.*, 1997). The *MUD2* synthetic lethal screen not only led to the identification of *BBP*, but also revealed a genetic link of *MUD2* with the U1 snRNP. The gene for the U1 snRNP associated Prp40 protein (Kao and Siliciano, 1996) was identified in this screen. Subsequent *in vitro* binding assays demonstrated, that BBP directly interacts with both Prp40p and Mud2p, but that Prp40p and Mud2p do not contact each other directly (Abovich and Rosbash, 1997). It is now widely agreed, that the three proteins form a cross-intron bridge through BBP contacting Prp40p in the U1 snRNP as well as the branchpoint sequence and Mud2p (see figure I.4). This cross-intron bridge brings the 5' end of the intron in proximity to the 3' end, and forms the structural basis for subsequent U2 snRNP addition. The importance of this arrangement is reflected in the conservation of the components involved: mammalian homologues of Prp40p (mPrp40p), BBP (SF1) and Mud2p (U2AF[65]) have been identified and it was shown that they act in an analogous way (Rain *et al.*, 1998 and Abovich and Rosbash, 1997 and references therein).
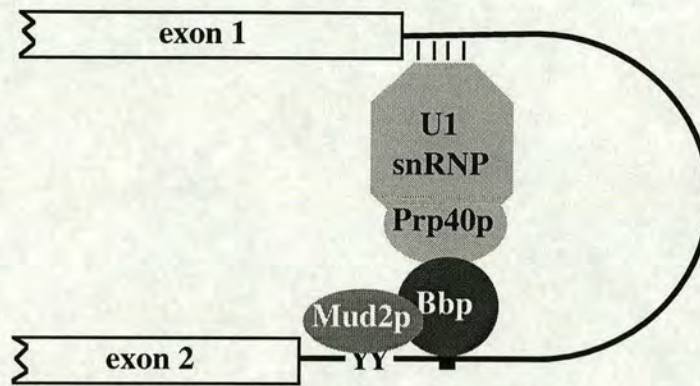
Figure I.4:
The "cross-intron bridge" model. The 5'- and the 3'-end of the intron are brought into proximity by the formation of a molecular bridge in commitment complex 2 (CC2). Exon sequences are represented as boxes, the intron as a line. Vertical lines indicate base pairing of the U1 snRNA with the 5' splice site. The U1 snRNP is depicted as octagon, the three other participating proteins Prp40p, Bbp and Mud2p are depicted as ovals/circles. Binding of Mud2p to the polypyrimidine tract (YY) and binding of Bbp to the branchpoint sequence (black square) is indicated.

I.3.2. Formation of the pre-spliceosome

Three essential U2 snRNP splicing factors, Prp9p, Prp11p and Prp21p (Legrain and Choulika, 1990; Schappert and Friesen, 1991; Arenas and Abelson, 1993) play an important role not only for the formation and the stability of the U2 snRNP but also for the formation of the pre-spliceosome, i.e. for U2 snRNP addition onto CC2. Many independent genetic as well as biochemical studies proved the existence of a complex of these three proteins, which is directly associated with the U2 snRNA (Legrain and Champon, 1993; Ruby and Abelson, 1993; Wells and Ares, 1994). Furthermore, two-hybrid studies established that Mud2p interacts with Prp11p and it was supposed that this interaction mediates the U2 snRNP addition onto CC2 (Abovich *et al.*, 1994). The importance of the Prp9p/Prp11p/Prp21p complex is reflected in its conservation: the structural similarity and the immunological cross-reactivity of the proteins suggest that they represent the homologues of the three subunits of the mammalian splicing factor SF3a (SF3a[60], SF3a[66] and SF3a[120]; Behrens *et al.*, 1993; Bennett and Reed, 1993;

Kraemer *et al.*, 1994, 1995). An important question that arises at this stage is, whether U2 snRNA association with the branch site sequence requires the replacement of BBP. Indeed, a recent study suggests that BBP as well as Mud2p are not present in pre-spliceosomes and thus, that they are displaced upon addition of the U2 snRNP (Rutz and Seraphin, 1999). U2 snRNP addition is the first ATP-dependent step during spliceosome formation. Two ATPases (and putative RNA helicases) therefore came to mind, which could be required to structurally reorganise the U2 snRNA prior to its association with the branchpoint region: Prp5p and Sub2p. Prp5p, at least *in vitro*, structurally alters the U2 snRNA, making its branchpoint recognition sequence more susceptible to targeted RNase H cleavage (O'Day *et al.*, 1996), suggesting that Prp5p action could increase the accessibility of the branchpoint recognition sequence for interaction with the intron. Sub2p is the putative yeast homologue of the mammalian splicing factor UAP56, which was demonstrated to be recruited to the spliceosome by U2AF[65], the mammalian homologue of Mud2p (Fleckner *et al.,* 1997). Thus, Sub2p, which was originally identified as suppressor of a cold-sensitive snRNP biogenesis mutant (Noble and Guthrie, 1996), could be at the site of action during pre-mRNA/U2 snRNA base pairing. RNA-helicases or unwindases would be expected to be particularly prone to cold, due to hyperstabilisation of their putative substrate.

I.3.3 Assembly and association of the U4/U6-U5 tri-snRNP

A number of proteins have been demonstrated to be required for the formation or the stability of the tri-snRNP. If they are mutated or depleted progression beyond the pre-spliceosome stage fails for that reason. Immunoprecipitation experiments revealed that both Prp3p and Prp4p are components of the U4/U6 di-snRNP and the U4/U6-U5 tri-snRNP (Banroques and Abelson, 1989; Anthony *et al.*, 1997). Later a direct physical interaction between these proteins was demonstrated by two-hybrid analysis and *in vitro* immunoprecipitation assays (Ayadi *et al.*, 1998). Using a heat-sensitive *prp3* mutant strain, it could be demonstrated that under non-permissive conditions the level of free U6 snRNA in the cells decreased, the level of free U4 snRNA increased and the levels of U4/U6 snRNA hybrids only slightly decreased. Looking at the

16

snRNPs, a depletion of free U6 snRNP and tri-snRNP but not of U4/U6 and U5 snRNP was observed. These results suggested a requirement of Prp3p for the formation of a stable U4/U6 di-snRNP and subsequent formation of the tri-snRNP (Anthony *et al.*, 1997). Glycerol gradient fractionation of snRNPs from heat-inactivated *prp4* mutant extracts revealed that as well as there being hardly any free U6 snRNP, there was also a strong decrease in U4/U6- and tri-snRNP levels, suggesting that Prp4p is required for U4/U6 di-snRNP stability (Galisson and Legrain, 1993).

Another protein, which is required for the formation of the tri-snRNP is Prp6p. Inactivation of a heat-sensitive mutant had no effect on U4/U6 di-snRNP levels but prevented tri-snRNP formation (Galisson and Legrain, 1993).

The 280 kDa U5 snRNP protein Prp8p was demonstrated to be essential not only for the assembly of the tri-snRNP, but also for the association of the tri-snRNP particle onto the pre-spliceosome: cell extracts depleted of Prp8p or extracts containing inactive Prp8p contained severely reduced levels of tri-snRNP particles. However, by adding anti-Prp8p antibodies, which did not affect tri-snRNP formation in extracts, spliceosome formation was blocked, suggesting an additional, more direct function for Prp8p in spliceosome assembly (Brown and Beggs, 1992).

Another two proteins have more recently been demonstrated not to be essential for U4/U6-U5 snRNP formation, but to be required for the recruitment of the tri-snRNP to the pre-spliceosome. In cell extracts prepared from cells containing a heat-sensitive mutant of the tri-snRNP-associated Prp31 protein, the levels of tri-snRNP were not significantly altered compared to wt cells, but the tri-snRNP failed to assemble onto the spliceosome, suggesting a direct role of Prp31p in the transformation of the pre-spliceosome into an active one (Weidenhammer *et al.*, 1997). The Clf1 protein (also named Syf3p) was shown to be a splicing factor required before the first transesterification reaction takes place. It interacted in two-hybrid analyses and in *in vitro* immunoprecipitation experiments with components of the commitment complex, Prp40p and Mud2p (Chung *et al.*, 1999). In extracts prepared from Clf1p-depleted cells, pre-spliceosomes formed, but tri-snRNP addition is severely decreased. The authors suggested that the protein might be involved in structural rearrangements within the pre-spliceosome which could be needed for tri-snRNP addition. Coimmunoprecipitation experiments with a hemagglutinin (HA) tagged version of the

protein, however, failed to detect any association with snRNAs. It was therefore proposed, that either the HA-tag might not be accessible to the antibodies when the protein is bound to an snRNP particle or the association with snRNP(s) is only weak. Alternatively, it was suggested, that the protein could be part of a larger, non-snRNP complex, which was supported by the fact that only approximately half of the HA-Clf1p could be precipitated from extracts and that the remainder was sufficient for splicing.

I.3.4 Toward an active spliceosome

Concurrently with and immediately after the association of the tri-snRNP onto the pre-spliceosome, major conformational rearrangements take place which involve fundamental changes in snRNA-snRNA and pre-mRNA-snRNA interactions. These are probably the most important events in bringing the reactive sites into the correct positioning within the catalytic centre(s) to initiate splicing (reviewed in Newman, 1994; Umen and Guthrie, 1995b; Nilsen, 1998). After binding of the tri-snRNP, the U4/U6 snRNA helix is disrupted and the U4 snRNP becomes destabilised from the spliceosome. The U6 snRNA changes its secondary structure and forms a novel helix by base pairing with the U2 snRNA, which is mutually exclusive from the U4/U6 snRNA-structure. The U1 snRNA/5' splice site interaction is also disrupted and replaced by new interactions of the U6 and U5 snRNAs with the 5' splice sites prior to step 1. The U6 snRNA base pairs with a short intronic sequence, while U5 associates with exon nucleotides adjacent to the 5' end of the intron.

Again a number of protein factors are known, which are present in or join the spliceosome at this particular stage and may promote the above mentioned rearrangements prior to the first transesterification. At least five proteins are implicated in supporting the U4/U6 snRNA dissociation, which is a prerequisite for all the subsequent steps.

The best candidate to date which could be directly responsible for U4/U6 snRNA unwinding is the 246 kDa DEIH-box ATPase and RNA helicase Brr2p (for reviews on DExD/H-box proteins see Hamm and Lamond, 1998; de la Cruz *et al.*, 1999). Brr2p is

unusual in that it contains two (instead of one for the other members of the family) conserved RNA helicase-like domains. Only the first one was shown to be essential for cell viability and pre-mRNA splicing (Kim and Rossi, 1999). Raghunathan and Guthrie (1998) immunopurified Brr2p in a native complex containing U1, U2, U5 and duplex U4/U6 snRNPs. They were able to show that upon addition of hydrolysable ATP to this complex the U4/U6 di-snRNP dissociated. However, if complexes were prepared from cells containing a mutation in the first helicase domain of *BRR2* (*brr2-1*), these rearrangements did not occur. Indeed, the investigation of the deproteinised RNA samples of this experiment by non-denaturing gel electrophoresis confirmed that in the wt preparations U4/U6 snRNAs were dissociated upon addition of ATP, but not in the preparations from the *brr2-1* mutant strain. These *in vitro* observations in yeast are strongly supported through investigations on the human homologue of Brr2p, the U5-200K protein (Laggerbauer *et al.*, 1998). The authors showed that partially purified U5 snRNPs containing the U5-200K protein are able to unwind U4/U6 duplex RNA in an ATP-dependent fashion. Depletion of U5-200K abolished unwinding activity. Furthermore they purified the U5-200K protein to near homogeneity and proved its ability to unwind the U4/U6 duplex *in vitro*. The first evidence for an U4/U6 unwinding function of Brr2p *in vivo* comes from a recent study in yeast: Kim and Rossi (1999) demonstrated that overexpression of dominant negative *BRR2* mutant alleles (the mutations were located in the first helicase domain) leads to an accumulation of duplex U4/U6 snRNAs in the cells, whereas the total levels of U4 and U6 snRNAs remained unchanged. Together with the *in vitro* data presented above, this provides good evidence for a direct role of Brr2p in U4/U6 snRNA unwinding in the cells.

Interestingly, certain thermosensitive mutants of the aforementioned tri-snRNP protein Prp4p do not prevent the formation of the tri-snRNP particle and its association with the pre-spliceosome, but nevertheless the formed spliceosome is not capable of splicing. In this case, as suggested by immunoprecipitation of spliceosomes from heat-treated mutant extracts with anti-Prp4p antibodies, the mutations affect the dissociation of the U4 snRNA and Prp4p from the spliceosome and therefore, necessary conformational rearrangements cannot take place to promote splicing (Ajadi *et al.*, 1997). Thus, Prp4p has been implicated in the release of U4 snRNA from the

spliceosome. Whether the protein directly effects U4/U6 snRNA unwinding remains to be determined. Prp38p, another tri-snRNP associated protein, seems to resemble the function of Prp4p quite well. Prp38p function is required for pre-mRNA splicing before the first transesterifcation reaction. In extracts made from a heat-sensitive mutant *prp38* strain, spliceosomes assembled, as shown by resolving the complexes in a native gel, but the extracts were incapable of splicing (Blanton *et al.*, 1992). Additional studies using a strain in which the expression of HA-tagged Prp38p could be conditionally regulated demonstrated the requirement of Prp38p for U4/U6 unwinding (Xie *et al.*, 1998). The authors used extracts either containing Prp38p or which were metabolically depleted of Prp38p and incubated these with a biotin-substituted pre-mRNA substrate. At various time points the assembled complexes were recovered by streptavidin chromatography and investigated for snRNA content. It could be demonstrated, that the levels of U4 snRNA in the complexes were significantly increased at later time points when Prp38p was absent. Native gel electrophoresis of the HA-Prp38p-containing and -depleted complexes, under conditions in which the U4/U6 hybrid can be resolved from the free species then confirmed that indeed most of the U4 snRNA at late time points was found associated with U6 snRNA in the Prp38p-depleted extracts. Since both tri-snRNP proteins, Prp38p as well as Prp4p, are unlikely to have RNA-unwindase activity themselves, it could be that they in some way support Brr2p function *in vivo*, maybe simply by providing the structural requirements for Brr2 positioning in the spliceosome.

An snRNA-free complex of at least seven proteins associated with the Prp19 protein is believed to join the spliceosome concurrently with or immediately after dissociation of the U4 snRNP (Tarn *et al.*, 1993a; 1993b and 1994) and therefore might play a role in U4 dissociation or at least in a step of spliceosome assembly that could be linked with U4 dissociation. Besides Prp19p, four other proteins of this complex have been characterised and were shown to associate with the spliceosome at the same time point: Cef1p (Tsai *et al.*, 1999), Snt309p (Chen *et al.*, 1998), Ntc20p and Ntc30p (Tsai *et al.*, unpublished results). The authors stabilise the normally very short-lived spliceosomal complex consisting of the U1/U2/U5/U4/U6 snRNPs (complex A2-1) by lowering the levels of ATP in splicing reactions. By using anti-Prp4p antibodies to immunoprecipitate spliceosomes from splicing reactions performed

at different ATP concentrations, they showed that the amount of pre-message that can be coprecipitated from the reactions increased with decreasing concentrations of ATP. Since Prp4p was previously demonstrated to associate with the spliceosome only in complex A2-1, this showed that the transition of the A2-1 complex into complex A1, containing only U2/U5/U6 snRNPs, could be slowed down by decreasing the levels of ATP.

Prp19 was shown previously, not to be present in the commitment complex or in the pre-spliceosome but to be essential for splicing before the first transesterification (Tarn *et al.*, 1993a). Since at high (2mM) ATP concentrations, coimmunoprecipitation of spliceosomes from splicing reactions using anti-Prp19p antibodies precipitates pre-mRNA, lariat-intron exon 2 and exon 1 intermediates as well as the excised intron, but at low ATP concentrations only very low levels of pre-mRNA were found in the precipitates, it was concluded that Prp19p must join the spliceosome concurrently with or after dissociation of U4 snRNP (Tarn *et al.*, 1993b). Snt309p, Cef1p, Ntc30p and Ntc20p exhibited the same pattern of spliceosome association as Prp19p. Furthermore, far-western blot analyses revealed a direct interaction of Snt309p and Cef1p with Prp19p and, notably, addition of anti-Cef1p antibodies prevents binding of Snt309p, Prp19p, Ntc20p and Ntc30p to the spliceosome, it was supposed that these proteins join the spliceosome as a preformed complex immediately prior to the first transesterification reaction (Tsai *et al.*, 1999). Prp19p and Cef1p are essential genes, which highlights the importance of this complex during spliceosome assembly.

Kuhn *et al.* (1999) reported that a mutation (U4-cs1) within the region of U4 snRNA that masks the ACAGA box in the U4/U6 complex conferred a reversible cold-sensitive block on splicing before the first transesterification reaction. The ACAGA box is the sequence within the U6 snRNA that base pairs with the 5' splice site prior to the first transesterification. The U4-cs1 mutation lies in a region of the U4 snRNA which lies opposite the ACAGA box in the U4/U6 RNA duplex and the mutation leads to hyperstabilisation of the helix by introducing additional Watson and Crick base pairings into the duplex. The authors show that in the cold not only U4/U6 snRNA unwinding but also U1 snRNA dissociation from the 5' splice site is prevented, suggesting that these processes are tightly coupled. In a search for trans-acting suppressor mutations a mutant allele of *PRP8* (*prp8-201*) was identified (Li and Brow,

21

1996). It was suggested that the normal function of Prp8p at this stage would be to downregulate U4/U6 snRNA duplex unwinding. Contact of Prp8p with the short U6 snRNA/5' splice site helix could induce conformational changes within Prp8p that would normally promote U4/U6 unwinding. In the mutant *prp8-201* these changes in Prp8p structure might be prevented. Alternatively, the function of Prp8p could be to negatively regulate a factor that unwinds the U4/U6 duplex, maybe Brr2p. For both models there is some experimental support and certainly the models are not mutually exclusive: i) It was demonstrated through crosslinking experiments using 4-thiouridine-modified pre-mRNA, that Prp8p indeed directly contacts the pre-mRNA substrate prior to the first transesterification reaction (Beggs *et al.*, 1995; Teigelkamp *et al.*, 1995). Contacts were observed with exonic nucleotides +1, +2 and + 8 upstream of the 5' splice site and thus Prp8p is in proximity to the 5' splice site/U6 snRNA duplex. ii) Prp8p was identified in a two-hybrid screen of a yeast genomic DNA library using Brr2p as bait and furthermore, the two proteins could be coimmunoprecipitated from cell extracts, making a contact of the two proteins during the splicing process likely (van Nues, unpublished results).

Prp8p not only contacts the pre-mRNA at this stage of the splicing reaction, but also the U5 snRNA (Dix *et al.*, 1998). Interestingly, the interaction of Prp8p with the U5 snRNA is strongest in a conserved sequence (loop 1) as shown by photo-crosslinking experiments. It was demonstrated previously that indeed loop 1 of the U5 snRNA directly contacts exon sequences adjacent to the 5' splice site prior to step 1 (and step 2, see later) (Newman and Norman, 1991, 1992). A role in stabilising the fragile U5 snRNA/pre-mRNA interaction has therefore been proposed for Prp8p. Surprisingly, the conserved loop 1 of the U5 snRNA was shown not to be essential for the first cleavage reaction, suggesting that other functions of the U5 snRNA and Prp8p at this stage are maybe more important or simply that this structural arrangement is needed at later stages. Maybe this also suggests that Prp8p is the primary factor to hold exon 1 in position.

The DEAD-box RNA helicase Prp28p is another good candidate to modulate the snRNA/pre-mRNA contacts at the 5' splice site prior to step 1. Staley and Guthrie (1999) generated reporter constructs in which the U1 snRNA/5' splice site duplex was hyperstabilised, by introducing additional nucleotides thereby extending the length of

the duplex. This hyperstabilisation led to a moderate decrease in splicing efficiency, which was exacerbated in the cold. Interestingly, altering the U6 snRNA sequence in a way that allowed extended base pairing of U6 snRNA to the 5' splice site increased splicing efficiency again, suggesting a competition of U1 and U6 snRNAs for binding to the 5' splice site. The authors then looked for mutants which could exacerbate the splicing defect by combining known splicing mutants together with the reporter constructs that conferred cold sensitivity due to the defect in 5' splice site cleavage. Among 15 splicing mutants tested, only a mutation in *PRP28* (*prp28-1*) enhanced the defect *in vivo*. Further evidence for a direct involvement of Prp28p in the switch of U6 snRNA for U1 snRNA at the 5' splice site was provided: Hyperstabilisation of the U1/5' splice site helix leads to a stall of spliceosomes, containing all spliceosomal snRNP particles, prior to step 1. These stalled spliceosomes could be activated by shifting the temperature to 35°C, whereupon U1 and U4 snRNAs are released. However, upon immunodepletion of Prp28p spliceosome formation is not prevented, stalled spliceosomes do form, but increasing the temperature does not lead to activation of the spliceosomes in this case. U1 and U4 snRNAs remain associated. Addition of recombinant Prp28p reverts this effect, supporting the idea of a requirement for Prp28p for the switch of basepairing of U6 snRNA for U1 snRNA at the 5' splice site prior to step 1.

Another putative RNA-helicase, the DEAH-box ATPase Prp2p is required for promoting the first transesterification reaction *in vivo* and *in vitro* (Lustig *et al.*, 1986; Plumpton *et al.*, 1994). It could be demonstrated that Prp2p joins the spliceosome transiently after association of the tri-snRNP and leaves before the second transesterification is initiated (King and Beggs, 1990). After preparation of fully assembled spliceosomes lacking Prp2p from ATP-depleted extracts, it could be shown that only the combined addition of wt Prp2p and ATP, but not one of the components alone was sufficient to promote splicing and dissociation of Prp2p, suggesting that ATP-dependent action of Prp2p is required prior to step 1. Teigelkamp *et al.* (1994) expressed and purified a dominant negative form of Prp2p, which contained a single amino acid exchange in a conserved motif, putatively required for RNA-helicase activity of the DEAD-box proteins. By adding this dominant negative Prp2 protein to splicing extracts before adding ATP and radiolabelled precursor-mRNA, splicing

complexes form but splicing becomes stalled prior to the first transesterification reaction, suggesting that RNA-helicase activity of Prp2p is required for initiation of the first step. The demonstration that Prp2p can be crosslinked to a spliceable pre-mRNA (mutations in the 5' splice site abolished the Prp2p crosslink) and that the crosslink is not dependent on the presence of a 3' splice site, suggested that the pre-mRNA might be the substrate for Prp2p and that its role could be to either resolve secondary structures within the pre-mRNA or to help displacing U1 snRNA from the pre-mRNA (Teigelkamp *et al.*, 1994). In a high copy-number suppressor screen with temperature sensitive *prp2* mutants, the *SPP2* gene was identified (Last *et al.*, 1987). Indeed, the initial analysis of the gene revealed that it is, like *PRP2*, required for splicing prior to step 1 *in vivo* and *in vitro* (Roy *et al.*, 1995). In extracts prepared from the *spp2*-mutant strain, spliceosomes assemble as normal, but are blocked prior to the first transesterification, showing that spliceosome assembly is not effected in the absence of Spp2p. When extracts containing these spliceosomes were depleted of ATP, it could be shown that again, as for Prp2p, only addition of both recombinant Spp2p and ATP could chase these extracts through splicing, indicating that Spp2p acts prior to or concomitantly with the ATP-hydrolysis required for step 1 of splicing. Using anti-Spp2 antibodies to precipitate spliceosomes, which were assembled in wt-, mutant-*spp2* or mutant-*prp2* extracts, it was demonstrated that Spp2 co-precipitated pre-mRNA only in the mutant-prp2 extracts, suggesting i) that it associates with the spliceosome prior to Prp2p action and ii) that under wt-conditions it is only very transiently associated with the spliceosome, leaving after ATP-hydrolysis by Prp2p, maybe associated with it. Two-hybrid analyses indeed suggests a physical interaction between the proteins (Roy *et al.*, 1995; Smith, 1999).

The protein actions presented above help to support the essential structural rearrangements of pre-mRNA and snRNAs. Immediately prior to step 1, a situation can be expected in which U6 snRNA base pairs with the 5' splice site as well as with the U2 snRNA. The U2 snRNA also base pairs to the branchpoint sequence and thus a situation could be imagined, where the branchpoint adenosine residue is in proximity to the 5' splice site to initiate the first transesterification (see figure I.5).
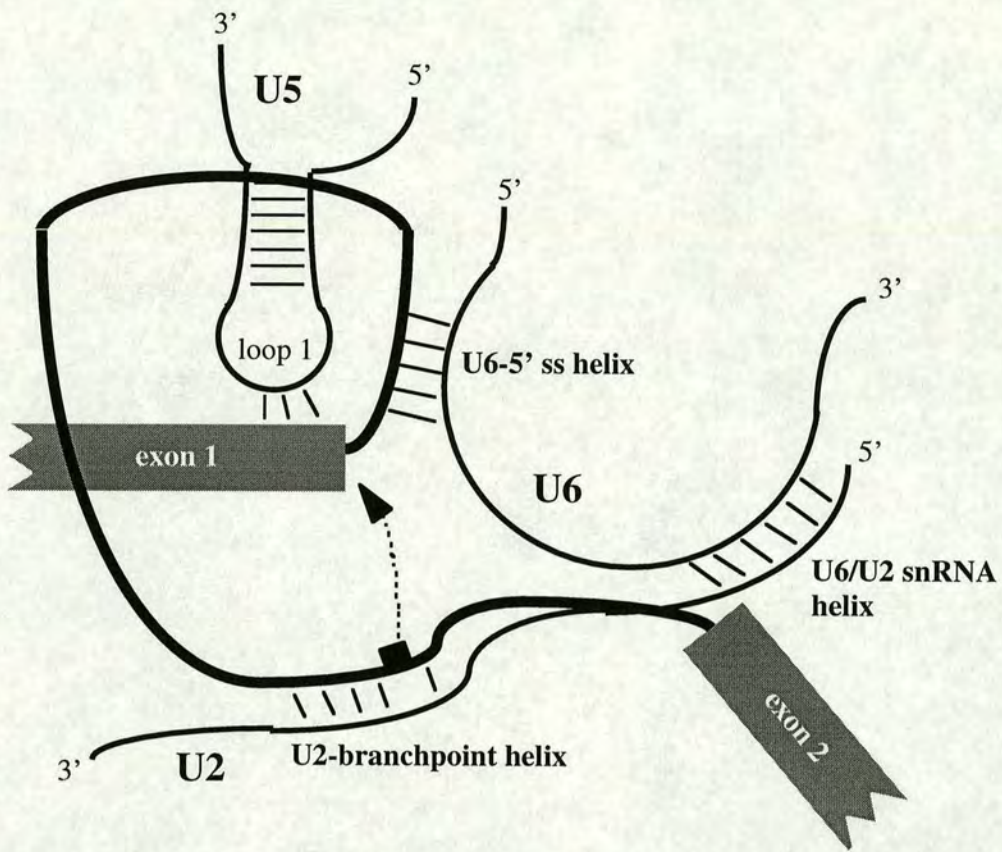
Figure I.5:
Schematic drawing of pre-mRNA/snRNA and snRNA/snRNA interactions immediately prior to step 1. Exon sequences are represented as dashed boxes, the intron as a fat line. U2, U5 and U6 snRNAs are depicted as thin lines. Short, thin lines indicate base pairing events: The U6 snRNA forms a helix with the 5' splice site as well as with the U2 snRNA (U6-5' ss helix and U6/U2 snRNA helix). U2 snRNA base pairs in addition with the branch sequence (U2-branchpoint helix). The U5 snRNA contacts the extreme 3' end of exon 1. The dotted arrow indicates the nucleophilic attack of the free 2'-hydroxyl group of the branchpoint adenosine (black square) on the phosphate group at the 5' splice site, which initiates the first step.

I.3.5 The second step of splicing

Although to date it is not clear whether the first and the second transesterifications happen in the same catalytic site, or whether two catalytic sites can be distinguished, at least some minimal rearrangements must be made in order to allow

the second transesterification to take place: i) the branchpoint must be displaced ii) the phosphodiesterbond at the 3' splice site must be brought into proximity to the free 3' hydroxyl group of exon 1 in order to allow the second nucleophilic attack for the ligation of the two exons to happen. Again, protein factors have been identified, which in some cases bind to the conserved cis-elements within the substrate pre-mRNA and which are specifically required for recognition and cleavage at the 3' splice site and subsequent linkage of the exons.

In addition to the 5' splice site and the branchpoint, which are also required for the first step, the two additional conserved elements within the pre-mRNA are important for the second step: the nucleotides PyAG at the 3' end of the intron and a polyuridine tract preceding the 3' splice site in most yeast introns. It has been demonstrated that mutations in any of the nucleotides PyAG inhibit the second transesterification reaction (Yijayraghavan *et al.*, 1986; Parker and Siliciano, 1993). The polyuridine tract preceding the 3' splice site seems to be required for 3' splice site selection. In constructs of introns with duplicated 3' splice sites in which only one is preceded by a polyuridine tract, it could be shown that this site will be preferentially used (Patterson and Guthrie, 1991).

Interesting is the fact that the 5' splice site as well as the branchpoint consensus sequence are necessary not only for the first, but also for the second transesterification. Three suggestions have been made to explain a requirement for these elements for the second step: first, in the majority of yeast introns, the branchpoint lies in proximity to the actual 3' cleavage site, which led to the model that the branch serves as a marker that facilitates 3' splice site recognition. Indeed, at least in mammalian *in vitro* splicing, there is some evidence for a scanning mechanism initiating at the branch structure (Smith *et al.*, 1989). Studies in yeast, however, using a number of reporter constructs which contain competing proximal and distal 3' splice sites, make a simple scanning mechanism unlikely (Luukkonen and Seraphin, 1997). Second, an inspection (or proofreading) mechanism has been proposed to enhance splicing fidelity, in which after the first transesterification, the branch structure is checked. Formation of the branch structure might be monitored by a protein binding (or trying to bind) to this structure, and if mistakes are recognised, the splicing intermediates become degraded.

26

A candidate monitoring protein for this mechanism could be Prp16p (Burgess and Guthrie, 1993). Mutations in the ATP-dependent RNA helicase Prp16p were isolated which allowed splicing of an aberrant lariat-intron exon 2 intermediate, that had formed due to a mutation in the branchpoint nucleotide. All seven isolated mutations within Prp16p were located in a conserved ATP-binding motif, and ATP-hydrolysis or binding was severely reduced in these mutants. The mutant lariats formed are rapidly turned over in wt cells, but are stabilised in the *prp16* mutant strains. The authors proposed the existence of a discard pathway under genetic control of Prp16p, through which in wt strains aberrantly formed lariat-intermediates are rapidly discarded. In *prp16* mutant strains, however, the decreased ATP-hydrolysis rate slows down the entry of these faulty intermediates into the discard pathway allowing more time for entry into a productive pathway. Third, evidence has been brought that nucleotides in the branch structure directly contact the conserved nucleotides PyAG at the 3' splice site (Parker and Siliciano, 1993).

Besides a number of proteins which will be discussed later, the U2, U6 and U5 snRNAs are essential transacting factors during the second step. Extensive mutational analyses have shown that both U2 and U6 snRNAs are required for the second transesterification to proceed. This is not surprising, because during the second step, both the U2 and U6 snRNAs maintain the base pairing interactions to each other and to the pre-mRNA, which have been established prior to the first step. It has to be noted, however, that certain nucleotides within the U2 and the U6 snRNA can be defined which are specifically required for just one of the transesterifications, indicating that the U2 and U6 snRNAs do not simply form a scaffold which is built before the first step and remains rather static in later stages of splicing, but that they are likely to take an active part in catalysis (for review see Umen and Guthrie, 1995b and references therein). Depletion of the U5 snRNA prior to step 2 blocks splicing before the second transesterification can take place (Winkelmann *et al.*, 1989). It was demonstrated that a four nucleotide sequence CUUU of loop 1 of the U5 snRNA contacts both exons adjacent to the 5' and the 3' splice sites (Newman and Norman, 1992; Newman *et al*, 1995). Photoactivated crosslinking using 4-thiouridine in the exon sequences revealed that the interaction of this loop can be detected with exon 1 prior to and during the first step as well as the second step. The interaction with exon 2 can only be detected after

completion of step 1 (Newman *et al.*, 1995). A model for the function of the U5 snRNA was proposed in which it fixes exon 1 in position after completion of the first step and then helps to bring the exons together through its interaction with exon 2. However, it is unlikely that both exons could be held together exclusively through these fragile base pairing events with the U5 snRNA. Additional factors are required.

One obvious candidate to help juxtapose the 3' end of exon 1 and the 5' end of exon 2 is Prp8p since, as mentioned above, it contacts the U5 snRNA as well as exon 1 already during the first step. Indeed, it could be shown by photo-crosslinking that after completion of the first step, Prp8p contacts the branch site and intron and exon nucleotides surrounding the 3' splice site (Teigelkamp *et al.*, 1995). One mutant of Prp8p (*prp8-101*) has been shown to be specifically defective in 3' splice site and polyuridine tract recognition and furthermore exhibits synthetic lethality with mutant alleles of a number of genes, encoding proteins that have been shown to be required for the second but not for the first transesterification such as Prp17p, Prp16p, Prp18p and Slu7p (Umen and Guthrie, 1995a), strongly supporting a role for Prp8p during the second step, in addition to its functions in spliceosome assembly and the first step.

A first indication that Slu7p and the Prp17p could be splicing factors came from their identification in a screen for mutants which are synthetic lethal with conditional alleles of the U5 snRNA (Frank *et al.*, 1992). The authors then provided evidence that these mutant alleles lead to a specific block of the second transesterification reaction *in vivo* and *in vitro*. Furthermore they showed synthetic lethality of mutant *slu7* and *prp17* genes with each other as well as with mutant alleles of *prp18* and *prp16*. In addition to *slu7* and *prp17*, synthetic lethality with certain mutations in the U5 snRNA loop 1 was also demonstrated for *prp8* and *prp18*, suggesting that the proteins are all involved in promotion of the second transesterification reaction. Indeed, evidence was provided that both Prp18p and Prp16p are bona fide second step splicing factors (Vijayraghavan and Abelson, 1990; Schwer and Guthrie, 1991; Horowitz and Abelson, 1993).

Prp16p was originally identified in a screen for suppressors of a branchpoint mutation (Burgess *et al.*, 1990), suggesting that Prp16p is involved in branchpoint recognition or binding. By now it is known that the DEAH-box protein has an RNA-dependent NTPase and ATP dependent RNA-helicase activity *in vitro* (Schwer and

Guthrie, 1991; 1992; Wang *et al.*, 1998). The use of non-hydrolysable ATPγS in the unwinding assay showed that indeed ATP-hydrolysis and not mere ATP-binding was required for unwinding of RNA-RNA duplexes, in a non-sequence specific manner. Extracts immunodepleted of Prp16p are blocked after the first transesterification has taken place. The addition of recombinant wt Prp16p can complement the second step block if ATP is present. However, if purified protein generated from the originally isolated *prp16* mutant was added to the extract prior to the wt Prp16p, complementation was abolished. Moreover, immunoprecipitation of spliceosomes from these extracts with anti-Prp16p antibodies demonstrated that the mutant Prp16p was stalled in the spliceosome, and therefore intermediates were efficiently precipitated, whereas from wt extracts Prp16p did not coprecipitate spliceosomes, showing its normally transient interaction with the spliceosome (Schwer and Guthrie, 1992). The investigation of the mutant revealed that the mutation resided in a conserved nucleotide binding domain and that the ATP-hydrolysis activity of the mutant was severely decreased. These data led to a model for Prp16p function, in which Prp16p binds to the branchpoint, hydrolyses ATP to promote structural changes within the pre-mRNA (and maybe within the spliceosomal complex) which are necessary for the second step to occur. Prp16p leaves the spliceosome immediately after this function.

To date, the ATP-dependent action of Prp16p is still the first action that can be defined after the first transesterification is completed. Umen and Guthrie (1995a) set out to identify the order of events during the second step in crosslinking studies in mutant and wt extracts by using a mutant substrate pre-mRNA, which contained a single amino acid exchange at the 3' splice site. By using this substrate the kinetics of the second step were slowed down, which could be demonstrated by enhanced UV-crosslinking of Prp8p to the mutant 3' splice site compared to the wt 3' splice site. The authors demonstrated that crosslinking of Prp16p to the 3' splice site is greatly reduced upon ATP-hydrolysis, as expected from previous observations (see above), but that crosslinking of Slu7p and Prp8p to the 3' splice site is greatly enhanced, suggesting that Prp16p action promotes binding of Slu7p and Prp8p to the 3' splice site. However, Slu7p and Prp8p binding is also detectable without the prior hydrolysis of ATP, but without conferring protection of the 3' splice site to oligonucleotide directed

RNAse H digestion. After ATP-hydrolysis, the 3' splice site is protected, suggesting strong binding of Slu7p and Prp8p upon conformational changes promoted by Prp16p action. When purified spliceosomes lacking Prp16p or Slu7p or both were generated and Prp16p, Slu7p and ATP were added back in a specified order it was shown that Slu7p action does not require ATP and furthermore it was confirmed that Slu7p acts after Prp16p function (Ansari and Schwer, 1995). Interestingly, although both Prp16p and Slu7p could be crosslinked to the 3' splice site, the removal of the corresponding region by RNase H digestion from preformed spliceosomes (generated by using wt pre-mRNA and extract immunodepleted of Prp16p) did not prevent the subsequent association of Prp16p and Slu7p with the spliceosome when they were added back. This suggested that at least the initial recruitment of the two factors might be mediated through protein-protein interactions. The same might be true for Prp17p and Prp18p, which could not be crosslinked to the 3' splice site at all (Umen and Guthrie, 1995a). In contrast to Slu7p and Prp16p, Prp17p and Prp18p are not essential for cell viability and depletion of Prp17p or Prp18p from extracts leads only to a partial block of splicing *in vitro* (Vijayraghavan *et al.*, 1989; Horowitz and Abelson, 1993). Prp17p acts before or concomitantly with an ATP-dependent reaction, whereas Prp18p action is ATP-independent (Horowitz and Abelson, 1993; Jones *et al.*, 1995). Thus, ordering the functions of the second step splicing factors strictly with respect to their ATP-requirement, it would place Prp16p/Prp17p action (ATP-dependent) prior to Slu7p/Prp18p action (ATP-independent). It was proposed that the non-essential factors Prp17p and Prp18p might modulate the function of their essential partners, Prp16p and Slu7p, respectively. This hypothesis is supported by the fact that overexpression of *PRP16* can suppress the phenotype of a *prp17* mutation and overexpression of *SLU7* can suppress a *prp18* mutation (Jones *et al.*, 1995).

Another DEAH-box protein was recently shown to be required for the second transesterification reaction: Prp22p (Schwer and Gross, 1998). When Prp22p was immunodepleted from whole cell extracts and splicing reactions were performed, the products of step 1, free exon 1 and lariat-intron exon 2 accumulated, demonstrating that Prp22p was required for the second step to proceed. Addition of recombinant Prp22p restored full splicing activity to the extracts. By using a mutant substrate pre-mRNA containing a nucleotide exchange at the 3' splice site to slow down the reaction

kinetics, the authors were able to demonstrate that, compared to wt extracts, in Prp22p-depleted extracts the 3' splice site was far more sensitive to oligonucleotide-directed RNase H cleavage, suggesting that Prp22p acts at the 3' splice site.

Earlier, Prp22p was identified as a factor required for the release of the mature message from the spliceosome (Company *et al.*, 1991). In extracts prepared from a *prp22* temperature-sensitive strain, heat inactivation prior to performance of splicing reactions led to an increased amount of excised intron as well as mRNA. By looking at the splicing complexes formed during the reactions it could be seen that in the heat inactivated extract, at late time points the spliced exons were found predominantly in a high molecular weight complex (the spliceosome), whereas in wt extracts the mRNA was shown to be present predominantly in lower molecular weight complexes. This suggested, that Prp22p was required to release the mature message from the spliceosome.

Interestingly, by using pre-mRNA substrates with different spacings between the 3' splice site and the branchpoint in Prp22p-depleted cell extracts, it was demonstrated that *in vitro*, Prp22p is dispensable for splicing of introns with 3' splice site-branchpoint distances of less than 21 nucleotides (Schwer and Gross, 1998). Slu7p and Prp18p exhibit a similar feature. Introns with short 3' splice site-branchpoint distances are efficiently spliced, but if the distance exceeds 11 nucleotides, Slu7p and Prp18p are absolutely required for progression through the second step (Brys and Schwer, 1996; Zhang and Schwer, 1997). The authors propose that for those introns in which the 3' cleavage site is close to the active centre of the spliceosome, which is supposed to reside at the branchpoint after step 1, the Prp16p-induced conformational change is sufficient to rearrange the active site for the second step. For 3' cleavage sites further away from the branch (and the active centre) Slu7, Prp18p and Prp22p could form a molecular bridge between the branch and the 3' splice site.

The general resemblance of the splicing pathway in mammals and yeast is very well illustrated by the conservation of the second step splicing factors. For all yeast second step splicing factors characterised to date, human homologues have been identified and for many, functional analogies have been described (Anderson *et al.*, 1989; Ohno and Shimura, 1996; Horowitz and Krainer, 1997; Zhou and Reed, 1998; Chua and Reed, 1999).

I.4   This thesis

In this work exhaustive two-hybrid screens of a yeast genomic DNA library have been performed to investigate protein-protein interactions within the spliceosome of *Saccharomyces cerevisiae*. The aim was, by using well characterised splicing proteins as baits, to reveal novel interactions between known splicing factors and possibly to identify hitherto unknown splicing proteins. Candidate proteins identified in the screens were investigated by generating null alleles of the genes, by constructing conditionally regulated alleles and/or by using the proteins in turn as baits in two-hybrid screens. Indeed, this procedure led to the identification of two novel pre-mRNA splicing factors, Prp45p and Prp46p, the initial characterisation of which is presented in this thesis. Both proteins are essential and evolutionarily conserved from yeast to man, which emphasises the importance of the proteins for the cell. Through the generation of strains allowing the conditional expression of *PRP45* and *PRP46*, it was demonstrated that both proteins are required for pre-mRNA splicing *in vivo*. By using a tagged-version of Prp45p in immunoprecipitation experiments, evidence was provided that the protein is an integral part of the spliceosome. When Prp45p was used as bait in a two-hybrid screen, it not only identified Prp46p, but also the proteins Syf1p and Syf3p, which have been recently demonstrated to be required not only for splicing but also for the progression of the cell cycle.

Chapter II

Materials and Methods

II.1 Materials

II.1.1 General reagents

II.1.1.1 Chemicals

Chemicals were purchased from the following sources, unless stated otherwise: Amersham, Fischer, Fisons, Melford Labs., National Diagnostics, Scotlab, Sigma.

II.1.1.2 Enzymes

Restriction enzymes, polymerases and other enzymes used in this work were purchased from the following companies, except where stated otherwise: Boehringer Mannheim, Gibco BRL, New England Biolabs (NEB), Pharmacia, Promega, Quiagen.

II.1.1.3 Growth reagents

Reagents for all growth media were purchased from: Beta Lab, Difco Laboratories, Oxoid or Sigma.

II.1.1.4 Antibiotics

All antibiotics were purchased from Duchefa (Netherlands).

II.1.2 Bacterial and yeast growth media

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

II.1.2.1 Bacterial media

<table>
<tr><td colspan="2" align="center">Table II.1.1
Bacterial media</td></tr>
<tr><td>Medium</td><td>Components</td></tr>
<tr><td>Luria-Broth (LB)</td><td>1% (w/v) Bacto-tryptone
0.5% (w/v) Yeast extract
0.5% (w/v) NaCl
adjust pH with 10 N NaOH to pH 7.2</td></tr>
<tr><td>SOC</td><td>2% (w/v) Bacto-tryptone
0.5% (w/v) Yeast extract
0.06% (w/v) NaCl
0.02% (w/v) KCl
0.1% (w/v) $MgSO_4$
0.4% (w/v) Glucose</td></tr>
<tr><td>10x M9 salts</td><td>6% (w/v) Na_2HPO_4
3% (w/v) KH_2PO_4
0.5% (w/v) NaCl
1% (w/v) NH_4Cl</td></tr>
<tr><td>M9 -L</td><td>10% -L (w/v) drop-out mix
0.2% (w/v) Glucose
10% (v/v) 10x M9 salts [a]
0.024% (w/v) $MgSO_4$ [a]
0.022% (w/v) $CaCl_2$ [a]</td></tr>
</table>

[a] added after autoclaving

II.1.2.2 Antibiotics

Antibiotics were added to liquid media immediatly prior to use. For solid medium, the antibiotic was added after autoclaving of the medium, when it was cooled down to approximately 50-60°C. Ampicillin and chloramphenicol were dissolved in 50% Ethanol to prepare stock solutions of a concentration of 100 mg/ml. These antibiotics were added to the media to give a final concentration of 100 µg/ml. Tetracycline was dissolved in 100% Ethanol to give a stock solution of 6 mg/ml. The final concentration in the media was 5 µg/ml. All antibiotics were stored at -20°C.

II.1.2.3 Yeast media

Table II.1.2 Yeast media	
Medium	Components
YPDA	1% (w/v) Yeast extract 2% (w/v) Bacto-peptone 2% (w/v) Glucose 0.003% (w/v) Adenine sulfate
YPD	as YPDA w/o Adenine sulfate
YMM[a]	0.67% (w/v) Yeast nitrogen base w/o amino acids 2% (w/v) Glucose
YMGRS[a]	0.67% (w/v) Yeast nitrogen base w/o amino acids 2% (w/v) Galactose 2% (w/v) Raffinose 2% (w/v) Sucrose
SPM (sporulation)	0.3% KAc 0.02% Raffinose

[a] note: To grow yeast, YMM or YMGRS was supplemented with complete drop-out powder (II.1.2.4). In this case, the medium was referred to as YMMsup or YMGRSsup. Selective media were supplemented with nutrients from complete drop-out powder but lacking the appropriate amino acid. In this case , the media will be referred to as YMM -X or YMGRS -X (X for any aminoacid), e.g. YMM -LW means, the medium contains all nutrients normally present in the complete drop-out powder mix, apart from leucine and tryptophan. Drop-out powder was added prior to autoclaving. For the preparation of solid medium 2% (w/v) agar was added prior to autoclaving.

II.1.2.4 Nutrients and supplements (Drop-out mix)

Drop-out powder was prepared by mixing 2g of each of the following nutrients: adenine, alanine, arginine, asparagine, aspartic acid, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, tryptophan, uracil, valine. If required 4g of leucine was included in the mix. The drop-out powder was intensively mixed and added to the media at 2% (w/v). To each 500 ml of medium containing drop-out powder, 325 µl of 10 N NaOH (for solid medium) or 250 µl of 10 N NaOH (for liquid medium) was added to adjust the pH.

II.1.3 Commonly used buffers

Table II.1.3 Commonly used buffers	
Buffer	Components
20x SSC	3 M NaCl 0.3 M Sodium citrate unsterilised
10x TE	100 mM Tris-HCl, pH 7.5 10 mM EDTA
10x TAE	0.4 M Tris-acetate. pH 7.5 20 mM EDTA unsterilised
10x TBE	0.9 M Tris-borate, pH 8.3 20 mM EDTA unsterilised
1x TBS	10 mM Tris-HCl, pH 7.5 150 mM NaCl unsterilised
1x TBS-TT	20 mM Tris-Cl, pH 7.5 500 mM NaCl 0.05% (v/v) Tween 20 0.2 % (v/v) Triton X-100 unsterilised

II.1.4 *Escherichia coli* strains

Bacterial strains used in this work are listed in table II.1.4. DH5αF' was used for all cloning procedures and propagation of plasmid DNA. MC1066 was used for propagation of plasmid DNA rescued from yeast cells (section II.3.2.4). BL21(DE3)pLysS was used for the expression of a recombinant yeast protein (section II.3.4.6).

<table>
<tr><td colspan="3" align="center">Table II.1.4
Escherichia coli strains</td></tr>
<tr><td>Strain</td><td>Genotype</td><td>Source</td></tr>
<tr><td>DH5αF'</td><td>F', φ80*dlac*ZΔM15, Δ(*lac*ZYA-*arg*F) U169, *deo*R, *rec*A1, *end*A1, *hsd*R17 (r_K^-, m_K^+), *sup*E44, λ⁻, *thi-1*, *gyr*A96, *rel*A1.</td><td>Gibco BRL.</td></tr>
<tr><td>MC1066</td><td>Δ(*lac*I POZYA)74, *gal*U, *gal* K, *Str*A^r, *leu*B6, *trp* C9830, *pyr*F74::Tn5(Kn^r), *hsd*R⁻.</td><td>P. Legrain</td></tr>
<tr><td>BL21(DE3) pLysS</td><td>F⁻, *omp*T, $hsdS_B$, (r_B^-, m_B^-), *gal*, *dcm* (DE3) pLysS</td><td>Novagen</td></tr>
</table>

II.1.5 *Saccharomyces cerevisiae* strains

Yeast strains used in this work are listed in table II.1.5.

Table II.1.5 *Saccharomyces cerevisiae* strains		
Strain	**Genotype**	**Source**
BMA38*	*MATa/α, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100*	B. Dujon Institute Pasteur
BMA38n	*MATα, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100*	This work
BMA64*	*MATa/α, his3−11,15, leu2-3,112, ura3-1, trp1Δ1, ade2-1*	F. Lacroute CNRS/CGM
BMA64n	*MATα, his3−11,15, leu2-3,112, ura3-1, trp1Δ1, ade2-1*	This work
CG1945	*MATa, ura3-52, his3Δ 200, ade2-101, lys2-801, trp1-901, leu2-3,112, lys2::GAL1$_{UAS}$-GAL1$_{TATA}$-HIS3, gal4-542, gal80-538, cyh^{r}2, URA3::GAL4$_{17-mers(x3)}$-CYC1$_{TATA}$-lacZ*	Clontech
Y187	*MATα, ura3-52, his3Δ 200, ade2-101, leu2-3,112, trp1-901, gal4 Δ, met$^-$, gal80 Δ, URA3::GAL4$_{17-mers(x3)}$-CYC1$_{TATA}$-lacZ*	Clontech
L40	*MATa, his3Δ200, trp1-901, leu2-3,112, ade2, lys2-801am, URA3::(lexA$_{op}$)$_8$-lacZ, LYS2::(lexA$_{op}$)$_4$-HIS3*	Hollenberg *et al.*, 1995

KY117	*MATa, his3Δ200, trp1Δ1, ade2, lys2-801, ura-3-52*	Chen and Struhl, 1985
KY118	*MATα, his3Δ200, trp1Δ1, ade2, lys2-801, ura-3-52*	Chen and Struhl, 1985
YMA156/1*	*MATa/α, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, SSY5/ssy5Δ::HIS3*	This work
YMA156/2	*MATa, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, ssy5Δ::HIS3*	This work
YMA44/1*	*MATa/α, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, YMR44c/ymr44cΔ::HIS3*	This work
YMA44/2	*MATa, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, ymr44cΔ::HIS3*	This work
YMA45/1*	*MATa/α, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, PRP45/HIS3-P_{gal1}-ProtA:PRP45*	This work
YMA45/2	*MATa, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, HIS3-P_{gal1}-ProtA:PRP45*	This work
YMA151KO1*	*MATa/α, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, PRP46/prp46 Δ::HIS3*	This work
YMA151/2	*MATα, his3−11,15, leu2-3,112, ura3-1, trp1Δ1, ade2-1, TRP1-P_{met3}-HA₂:PRP46*	This work

* all diploid strains are isogenic for the auxotrophic markers described

YMA45/1*	*MATa/α, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, PRP45/HIS3-P$_{gal1}$-ProtA:PRP45*	This work
YMA45/2	*MATa, his3Δ200, leu2-3,112, ura3-1, trp1Δ1, ade2-1, can1-100, HIS3-P$_{gal1}$-ProtA:PRP45*	This work
YMA151/2	*MATα, his3−11,15, leu2-3,112, ura3-1, trp1Δ1, ade2-1, TRP1-P$_{met3}$-HA$_2$:PRP46*	This work

II.1.6 Oligonucleotides

Oligonucleotides used in this work were purchased from Bioline Ltd. (London), Genosys Biotechnologies Ltd. (Cambridge) or Oswel DNA Service (Southampton).

	Table II.1.6 Oligonucleotides	
Name	**Sequence (5'-3')**	**Description (template)**
N3027	TCATCGGAAGAGAGTAG	pAS2ΔΔ for.
P5148	ATAAATCATAAGAAATTCGC	pAS2ΔΔ rev.
W2248	CTTCGTCAGCAGAGCTTC	pBTM forward
lexA reverse	TTTTAAAACCTAAGAGTCAC	pBTM reverse
T3785	GAAATTGAGATGGTGCACGATGCAC	pACTII 3'PCR
T3786	CGCGTTTGGAATCACTACAGGGATG	pACTII 5'PCR
JC90	GGCTTACCCATACGATGTTC	pACTII 5'PCR
U3exon2	CCAAGTTGGATTCAGTGGCTC	U3 snoRNA
U1	CAATGACTTCAATGAACAATTAT	U1 snRNA
T7-1	TAATACGACTCACTATAGGG	pET19b
156-1	CTTTTTGCATGTACATAGTACTGGTGTAAA CTCGATATACCGCTCTTGGCCTCCTCTAG	Δssy5
156-2	GCAAATCATCCATCTAGTTGTGGATCAATG TCCCATTGAATTTCGTTCAGAATGACACG	Δssy5
44ko1	TTGTTAACTACATTTTTCAGAACGGCGTGT CATTCTCCGATACTCTTGGCCTCCTCTAG	Δymr44w
44ko2	TATTGTTCAAAAGCAGAGTACTACAACTGC AATAGCAACAGGTCGTTCAGAATGACACG	Δymr44w
PFun-1	TTACCTTAACGTATTATTGTAATTCTTCAC GAATTTGATTCTCTTGGCCTCCTCTAGT	$P_{GAL1}::PRP45$
PFun-2	CTTGAGAATGTTTTGGAGGTGGTAGTCTGT TACTAAACATATTCGCGTCTACTTTCGG	$P_{GAL1}::PRP45$
Func1	TATGAATTCATGTTTAGTAACAGAC	PRP45 cloning
Func2	TATGTCGACCTAGGCGCCATAGTTATCC	PRP45 cloning
275A	ATGAGCTCTGAGTGAACTGCTTGCC	YOR275c
275B	TATAGTCGACTTTGAGGGATTCGGC	YOR275c
44A	ATGAGCTCCTGAAGCGATATTCCAG	YMR44w
44B	ATATGTCGACATCAACTGCAATAGC	YMR44w
7A	ATCCATGGATGATGACTGGGACTCC	SLU7

7B	AAGAATTCCGTTCAAGTAAGCAGCC	*SLU7*
FCR-1	TATGGATCCGAGCAATAGTTAAGCCC	*PRP45cr*
FCR-2	TATGGATCCTCGGATACGTCTCTTCC	*PRP45cr*
PLKO-1	GAGGATGCAGACACTGTGTTACATGGAGATTAGT GAGACTCTTGGCCTCCTCTAG	*Δprp46*
PLKO-2	CACGTATACAGGGTACGTACTTTTTCCATCTACT CCCATCGTTCAGAATGACACG	*Δprp46*
151MetA	CCACACAAATCCACGATGACCTAAGAACATTCGT TCGCTTAAATCCTTAAATAAATACTACTC	*P MET3::PRP46*
151MetB	TGTCTACATCTCCTAAATTTTCGACTTTGTGATC ATTTCCGTCCATACGAGCTCCAGCGTAATCTGGA A	*P MET3::PRP46*
151-PR2	GTATAAAGCCGAAGTCC	*PRP46 3'UTR*
Met3	TTTAGCTTGTGATCTC	*P MET3*
PGK	ACCGTTTGGTCTACCCAAGTGAGAAGCCAAGACA	*PGK probe*
RP28A	TCGTACTGATGCTCCATTC	*RP28 PCR*
RP28B	TGAAACCCTTAGATCTTC	*RP28 PCR*
G8102	CACGCCTTCCGCGCCGT	U1 snRNA
G8103	CTACACTTGATCTAAGCCAAAAG	U2 snRNA
483A	CCGTGCATAAGGAT	U4 snRNA
485A	AATATGGCAAGCCC	U5 snRNA
Taq6A	TC(A/T)TCTCTGTATTG	U6 snRNA

II.1.7 Plasmids

The plasmids used in this work are described in tables II.1.7 and II.1.8.

<table>
<tr><td colspan="3" align="center">Table II.1.7
General plasmids</td></tr>
<tr><td>Plasmid</td><td align="center">Features</td><td>Reference</td></tr>
<tr><td>pAS2ΔΔ</td><td>Two-hybrid bait plasmid. *GAL4* DNA-binding domain (GBD) fusion expression vector: Multiple cloning site, Amp^R, *col*E1 ori, P_{ADH1}, *GAL4* DNA binding domain sequence, *ADH1* transcriptional terminator, 2μ, *TRP1*.</td><td>Fromont-Racine *et al.*, 1997</td></tr>
<tr><td>pAS2ΔΔBg</td><td>Two-hybrid bait plasmid. Frame shifted version of pAS2ΔΔ in (-1) reading frame.</td><td>A. Colley, this laboratory</td></tr>
</table>

pBTM116	Two-hybrid bait plasmid. LexA DNA-binding domain fusion expression vector: Multiple cloning site, Amp^R, colE1 ori, P_{ADH1}, LexA sequence, ADH1 transcriptional terminator, 2μ, TRP1.	S. Fields, S.U.N.Y., Stony Brook
pBTM116(-1)	Two-hybrid bait plasmid. Frame shifted version of pBTM116 in (-1) reading frame	A. Mayes, this laboratory
pGBT9	Two-hybrid bait plasmid. Multiple cloning site, Amp^R, colE1 ori, P_{ADH1}, GAL4 DNA-binding domain sequence, ADH1 transcriptional terminator, 2μ, TRP1.	Clontech Laboratories, Inc.
pACTIIStop	Two-hybrid prey plasmid vector. GAL4 activation domain fusion shuttle and expression vector: Multiple cloning site, Amp^R, colE1 ori, P_{ADH1}, GAL4 activation domain sequence, HA-epitope sequence, ADH1 transcriptional terminator, 2μ, LEU2.	Fromont-Racine et al., 1997
pTL27	Yeast-E. coli shuttle vector: Multiple cloning site, Amp^R, colE1 ori, P_{GAL1}, 2x Protein A-epitope sequence, CEN6, ARSH4, HIS3.	Lafontaine and Tollervey, 1996
pNOPPATA1L	Yeast-E. coli shuttle vector: Multiple cloning site, Amp^R, colE1 ori, P_{NOP1}, 2x Protein A-epitope sequence, TEV protease cleavage site sequence, ADH1 transcriptional terminator, ARS1, CEN4, LEU2.	Klaus Hellmuth
pUC19-55HA$_2$	pUC19 based E. coli cloning vector. Modified: P_{MET3}, 2x HA-epitope sequence, TRP1.	R. van Nues, this laboratory
pET19b	E. coli expression vector. Cloning site, T7 promotor sequence, His-tag coding sequence, Enterokinase cleavage site sequence, T7 transcription termination sequence, Amp^R, lacI, colE1 ori.	Novagen
p283	Modified pGEM1. Contains ACT1 coding sequence under regulation of the T7 promotor.	O'Keefe et al., 1996
YIp1	Yeast integrative vector : Amp^R, HIS3.	Struhl et al., 1979

Table II.1.8
Modified plasmids

Plasmid	Features	Reference
pASSlu7	GBD::Slu7p encoding two-hybrid bait vector. Modified pAS2ΔΔ: Full length SLU7 ORF fused to GAL4 DNA-binding domain (GBD).	I. Dix, this laboratory

pASPrp18	GBD::Prp18p encoding two-hybrid bait vector. Modified pAS2ΔΔ: Full length *PRP18* ORF fused to *GAL4* DNA-binding domain.	I. Dix, this laboratory
pBSKS22	Modified pBluescript II KS+: contains *Sac*I, *Kpn*I fragment of *PRP22* locus cloned into multiple cloning site via *Sac*I and *Kpn*I. *Nde*I restriction site is introduced into ORF start: CAT**ATG**.	Beate Schwer (New York)
pMA22	GBD::Prp22p encoding two-hybrid bait vector. Modified pAS2ΔΔ: pBSKS22 was cut with *Kpn*I and blunt ends were generated by *T4* polymerase treatment. After *Nde*I restriction, the *PRP22*-containing fragment was isolated and ligated into pAS2ΔΔ cut with *Nde*I, *Sma*I. The identity of the insert and the frame was checked by sequencing.	This work
pMA22s	GBD::Prp22p(479-826) encoding two-hybrid bait vector. Modified pAS2ΔΔ: the prey plasmid insert from clone 7-III-5 identified in two-hybrid screen Slu7-III was isolated via *Bam*HI restriction and ligated into *Bam*HI linearised and dephosphorylated pAS2ΔΔ. Correct orientation and frame were verified by sequencing.	This work
pMA22l	GBD::Prp22p(142-747) encoding two-hybrid bait vector. Modified pAS2ΔΔ: the prey plasmid insert from clone 7-III-9 identified in two-hybrid screen Slu7-III was isolated via *Bam*HI restriction and ligated into *Bam*HI linearised pAS2ΔΔ. Correct orientation and frame were verified by sequencing.	This work
pMA275N	Modified pAS2ΔΔBg: An n-terminal 400 bp fragment of ORF *YOR275c* was PCR-amplified from yeast genomic DNA using oligonucleotides 275A and 275B. The PCR product was cut *Ecl*136II and *Sal*I and ligated into *Sma*I, *Sal*I cut pAS2ΔΔBg. Frame and integrity of the cloned PCR product was checked by sequencing.	This work
pMA275NC	Modified pMA275N: An 800 bp fragment containing the c-terminus of ORF *YOR275c* was isolated from p7-IV-75 via *Sal*I, *Nsi*I restriction. The fragment was ligated into *Sal*I, *Pst*I digested pMA275N. The identity of the cloned fragment was verified by sequencing.	This work
pMA275	GBD::Yor275p encoding two-hybrid bait vector. Modified pMA275NC: pMA275NC was linearised by *Sal*I restriction digest, gel purified and dephosphorylated. The linearised vector was then transformed into yeast strain BMA38, where the *YOR275c* ORF was generated via gap-repair. The repair event was verified by sequencing.	This work

pMA45	LexA::Prp45p encoding two-hybrid bait vector. Modified pBTM116: The *PRP45* ORF was PCR-amplified from yeast genomic DNA using oligonucleotides Func1 and Func2. The PCR product was gel-purified, cut with *Eco*RI, *Sal*I and ligated into *Eco*RI, *Sal*I restricted pBTM116. The *PRP45* coding sequence was verified by sequencing.	This work
pMA45cr	LexA::Prp45p(259-286) encoding two-hybrid bait vector. Modified pBTM116: The sequence encoding amino acids 259-286 of Prp45p was PCR-amplified from plasmid p22-1 using oligonucleotides FCR-1 and FCR-2 and the fragment was gel-purified and digested with *Bam*HI. It was ligated into pBTM116, *Bam*HI linearised and dephosphorylated. Orientation and integrity of the insert were verified by sequencing.	This work
pMA44AS	GBD::Ymr44p encoding two-hybrid bait vector. Modified pAS2ΔΔBg: The *YMR44c* ORF was PCR-amplified from yeast genomic DNA using oligonucleotides 44A and 44B. The PCR product was gel-purified, restricted with *Ecl*136II, *Sal*I and ligated into *Sma*I, *Sal*I digested pAS2ΔΔBg. The *YMR44c* coding sequence was verified by sequencing.	This work
pMA44BTM	LexA::Ymr44p encoding two-hybrid bait vector. Modified pBTM116(-1). Generation of *YMR44c* ORF as described for pMA44AS. The PCR fragment was ligated into *Sma*I, *Sal*I cut pBTM116(-1).	This work
pCR117	LexA::Syf3p encoding two-hybrid bait vector. Modified pBTM116. The *SYF3* ORF is ligated into the multiple cloning site of pBTM116 fused to the *LexA* coding sequence.	I. Dix, C. Russell, this laboratory
pGBT9/SNW1	GBD::SNW1 encoding two-hybrid bait vector. Modified pGBT9: The SNW1 coding sequence is fused to the GAL4 DNA-binding domain sequence in the multiple cloning site of pGBT9.	Mike Hayman, S.U.N.Y., Stony Brook
pETMA45	Modified pET19b (Novagen): The *PRP45* ORF was isolated from pMA45 via *Eco*RI, *Sal*I restriction. 3' recessing termini were filled in by Klenow treatment. The fragment was then ligated into pET19b, which had been linearised by *Nde*I restriction and blunted by Klenow treatment.	This work
pMA7c	GAD::Slu7p(164-257) producing two-hybrid prey plasmid. Modified pACTIIStop: The sequence of the *SLU7* ORF encoding amino acids 164-257 was PCR-amplified using oligonucleotides 7A and 7B. The PCR product was gel-purified and cut with *Nco*I, *Eco*RI and ligated into pACTIIStop, which had been cut *Nco*I, *Eco*RI. The insert was sequenced.	This work

p22	GAD::Prp22p encoding two-hybrid prey vector. pMA22 was linearised with *Nde*I, blunt ends were created by Klenow treatment, the fragment was cut with *Sal*I and the PRP22 containing fragment isolated. Subsequently it was ligated into pACTIIStop previously digested with *Sma*I and *Xho*I.	This work
p7-III-5	Two-hybrid library prey plasmid isolated in screen Slu7-III. Encodes fusion protein GAD::Prp22p(479-826).	This work
p7-III-9	Two-hybrid library prey plasmid isolated in screen Slu7-III. Encodes fusion protein GAD::Prp22p(142-747).	This work
p7-IV-48	Two-hybrid library prey plasmid isolated in screen Slu7-IV. Encodes fusion protein GAD::Prp18p(101-219).	This work
p7-IV-27	Two-hybrid library prey plasmid isolated in screen Slu7-IV. Encodes fusion protein GAD::Ssy5p(129-560).	This work
p7-IV-75	Two-hybrid library prey plasmid isolated in screen Slu7-IV. Encodes fusion protein GAD:Yor275p (108-661).	This work
pM18-4	Two-hybrid library prey plasmid isolated screen M18. Encodes fusion protein GAD::Slu7p(1-279).	This work
pM18-5c	Two-hybrid library prey plasmid isolated in screen M18. Encodes fusion protein GAD::Ixr1p(341-535)::Slu7(161-249).	This work
pM18-45	Two-hybrid library prey plasmid isolated in screen M18. Encodes fusion protein GAD::Slu7p(165-382).	This work
p22-1	Library prey plasmid isolated in Prp22p two-hybrid screen. Encodes fusion protein GAD::Prp45p(1-350).	This work
p22-10	Library prey plasmid isolated in Prp22p two-hybrid screen. Encodes fusion protein GAD::Prp45p(9-290).	This work
p22-13	Library prey plasmid isolated in Prp22p two-hybrid screen. Encodes fusion protein GAD::Prp45p(194-379).	This work
p22-19	Library prey plasmid isolated in Prp22p two-hybrid screen. Encodes fusion protein GAD::Prp45p(212-379).	This work
p22-27	Library prey plasmid isolated in Prp22p two-hybrid screen. Encodes fusion protein GAD::Prp45p(262-359).	This work
p22-12	Library prey plasmid isolated in Prp22p two-hybrid screen. Encodes fusion protein GAD::Syf3p(592-687).	This work

II.1.8 Antisera

The antisera used in this work are listed in Table II.1.9.

Table II.1.9 Antisera		
Antibody	Description	Source
Anti-Gal4 DNA-binding domain	Rabbit polyclonal antibodies. 1:1000 dilution for western blots.	Santa Cruz Biotechnology
Anti-Prp8p	Rabbit polyclonal antibodies raised against a 35 amino acid peptide of the n-terminal region of Prp8p. 15 μl used per immunoprecipitation.	This laboratory
Anti-His	mouse monoclonal anti-Penta-His antibody (0.2mg/ml), 1:3500 dilution for western blots.	Quiagen
Anti-rabbit IgG-HRP	Anti-rabbit IgG horse radish peroxidase linked whole antibody (from donkey). 1:5000 dilution for western blots .	Amersham
Anti-mouse IgG-HRP	Anti-rabbit IgG horse radish peroxidase linked whole antibody (from sheep). 1:5000 dilution for western blots .	Amersham
Anti-rabbit IgG-AP	Anti-rabbit IgG (Fc) alkaline phosphatase conjugate. 1:5000 for western blots.	Promega

II.2 Microbiological methods

II.2.1 Growth of strains

II.2.1.1 Growth of bacteria

E. coli strains were routinely grown at 37°C in rich LB medium (table II.1.1). To maintain selection for plasmid DNA, transformed bacteria were grown in rich medium containing the appropriate antibiotic(s) (section II.1.2.2).

II.2.1.2 Growth of yeast

Yeast strains were routinely grown at 30°C in YPDA or YMGRSsup (Table II.1.2). To maintain selection for plasmid DNA or to select transformants for integration of a reporter gene into the genome cells were grown at 30°C in YMMsup or YMGRSsup medium lacking the appropriate amino acid in the drop-out powder (YMM -X or YMGRS -X; section II.1.2.3, table II.1.2).

II.2.2 Preservation of strains

II.2.2.1 Preservation of bacteria

E. coli strains were stored for up to two weeks on solid medium at 4°C. Strains were stored indefinitely at -70°C in 15% glycerol. A stationary culture of the strain to be stored was grown in the appropriate medium. Eight hundred microlitres of culture was mixed with an equal volume of 30% glycerol and snap-frozen on dry ice.

II.2.2.2 Preservation of yeast

Yeast strains were stored for up to four weeks on solid medium at 4°C. Strains were stored indefinitely at -70°C in 25% glycerol. A culture of the strain to be stored was grown to mid-logarithmic phase in appropriate medium. Six hundred and forty microlitres of culture was mixed with 360 μl of 70% glycerol and was immediately frozen at -70°C.

II.2.3 Sporulation of yeast

II.2.3.1 Growth of diploids and sporulation

Diploid yeast cells of the strain to be sporulated were grown overnight at 30°C in YPDA medium. 1 ml of culture was transferred to a microcentrifuge tube and cells were sedimented in a benchtop centrifuge at 6000 rpm for 15 seconds. The supernate was removed and the cells were resuspended in 3 ml of minimal sporulation medium (SPM, table II.1.2). In a test tube, placed on a rotating wheel, the cells were incubated at 23°C for 5 to 7 days. The cells were examined microscopically to determine whether sporulation and tetrad formation had occurred.

II.2.3.2 Tetrad dissection

Upon successful sporulation and tetrad formation, 200 μl of cells were sedimented by centrifugation and resuspended in 200 μl of sterile, distilled water. After addition of 10 μl β-glucuronidase (10U/μl) the suspension was mixed gently and incubated at room temperature for 20-60 minutes. The suspension was serially diluted and spread in one line onto YPD or YMGRSsup (table II.1.2) agar plates. Tetrads were dissected using a Singer MSM system micromanipulator. After dissection, the spores were incubated at 23°C for 4 to 5 days.

II.2.3.3 Determination of mating type

Haploid progeny (auxotroph for leucine, prototroph for lysine) were propagated on solid media (table II.1.2). KY117 and K118 cells (Leu+, Lys-) were simultaneously grown on YPDA (table II.1.2) and then replica plated onto the cells of undetermined mating type on YPDA. The plates were incubated overnight and replica plated onto YMM -LK medium. Only formed diploid cells were able to grow. The mating type of the haploid progeny could thus be determined because only *MATa* cells can mate with KY118 (*MATα*), whereas only *MATα* cells are able to mate with KY117 (*MATa*) cells.

II.2.4 Transformation of *E. coli*

II.2.4.1 Preparation of electro-competent cells

5 ml of LB liquid medium (table II.1.1) was inoculated with a single *E. coli* colony and incubated overnight at 37°C. An aliquot of this culture was used to inoculate 500 ml of LB liquid medium to an optical density at 600 nm (OD_{600}) of 0.1 units. The culture was grown to an OD_{600} of 0.6-0.8 units. The culture was then placed on ice for 30 min. The cells were transferred to pre-chilled centrifuge beakers and were centrifuged for 15 min at 4000 rpm (4°C, Beckman 10.500 rotor). The cells were washed twice in 250 ml of ice-cold sterile 10% glycerol and sedimented as before. Finally, the cells were resuspended carefully in 1 ml ice-cold GYT (10% (v/v) glycerol, 0.125% (w/v) yeast extract and 0.25% (w/v) tryptone), aliquoted to 40 μl and frozen for at least 30 minutes at -70°C. The competent cells can be used for several months, if stored at -70°C.

II.2.4.2 Transformation of electro-competent cells

Forty microlitres of electro-competent cells were thawed on ice and mixed with 2 μl (10-100 ng) transforming plasmid DNA. The cells were transferred to an electroporation cuvette (0.2 cm electrode gap) on ice. Residual ice on the outside of the cuvette was wiped off with a paper tissue to ensure the cuvette was dry prior to electroporation. The cuvette was tapped carefully several times onto the bench to remove trapped air in between the electrodes. Electroporation was performed using a Biorad Gene Pulser II set at 200 ohms, 25 μF and 2.5 Kvolts. Immediately after electroporation, 1 ml of SOC medium (table II.1.1) was added and the cells were transferred to a fresh microcentrifuge tube. The cells were allowed recovery by incubation on a rotating wheel for 45 min at 37°C. Finally, the cells were plated on solid medium supplemented with the appropriate antibiotic to maintain selection for the transformed plasmid. The plates were incubated overnight at 37°C.

II.2.5 Transformation of yeast

Yeast cells were transformed using the method described by Gietz et al. (1992). A 10 ml overnight culture of yeast cells to be transformed was grown to a cell density of 1-2 x 10^7 cells/ml at 30°C. The cells were then diluted into 50 ml of pre-warmed medium to a density of approximately 2 x 10^6 cells/ml (approximately $OD_{600} = 0.1$) and regrown to a density of approx. 2 x 10^7 cells/ml ($OD_{600} = 0.7$) at 30°C. The cells were harvested by centrifugation at 3,500 rpm (Mistral 1000 centrifuge) for 3 minutes at room temperature, washed with 25 ml of sterile destilled water and harvested again as above. The cell pellet was then resuspended in 1 ml of 100 mM LiAc and transferred to a microcentrifuge tube. The cells were sedimented by centrifugation for 15 seconds at 13,000 rpm in a benchtop centrifuge. The supernate was removed and the cells were resuspended to a final volume of 500 μl (normally yields a cell density of 2 x 10^9 cells/ml) in 100 mM LiAc and subsequently split into aliqouts of 50 μl. The cells were sedimented as above and the supernate was removed. To each sample was added in the order listed: 240 μl 50% PEG_{3350} (w/v), filtersterilised, 36 μl 1M LiAc,

25 μl single stranded salmon sperm carrier DNA (2 mg/ml; boiled for 5 minutes and put on ice prior to use) and 5-25 μl DNA (0.1-2 μg/μl). Each tube was vortexed vigorously for 1 minute and incubated for 30 minutes at 30°C. The cells were then heat-shocked by transferring the tubes into a water bath at 42°C for 15 to 20 minutes. Subsequently, the cells were harvested by centrifugation (6,000-8,000 rpm, 15 seconds). The supernate was removed and the cells were resuspended in 1 ml sterile water by pipetting up and down gently. Finally, 250 μl aliquots were spread onto the appropriate selective solid media (table II.1.2). The plates were incubated at 30°C for 2-4 days.

II.2.6 ORF replacements and construction of conditionally regulated genes in yeast

II.2.6.1 ORF replacements

Replacements of open reading frames in yeast were performed using the method of Baudin *et al.* (1993). A linear DNA fragment containing the *HIS3* gene as marker flanked by 35-40 bp of sequence identical to the regions immediately upstream and downstream of the ORF to be replaced was generated by PCR (section II.3.2.8.1). The PCR was performed using genomic DNA of strain BMA38 (table II.1.5) as template. The primers (table II.1.6) used were typically about 65 bp in length, 45 bp for the homologous flanking regions and 15-20 bp for priming the amplification of the *HIS3* marker. The linear DNA fragment was transformed (section II.2.5) into yeast strain BMA38. Transformants were selected on solid YMM -H medium (table II.1.2). Histidine prototrophic diploids were streaked onto YMM -H medium for colony purification and were subsequently investigated for correct integration of the replacement cassette into the target locus by southern blotting (section II.3.2.17) or PCR-analysis (section II.3.2.8). Haploid cells with the replaced target locus were then produced by sporulation (section II.2.3.1) and subsequent tetrad analysis (section II.2.3.2).

II.2.6.2 Construction of conditionally regulated genes

Conditionally regulated genes were generated using the same principal method that was used for ORF replacements (section II.2.6.1). To construct a conditionally regulated allele of a gene, however, an integration cassette was PCR-generated (section II.3.2.8) in which the flanking regions were homologous to regions of the extreme 5' end of the target ORF and to a region approximately 100-200 nucleotides upstream of the target ORF. This should allow replacement of the native promotor of the gene (or at least partial deletion of it). The integration cassette was chosen to contain an auxotrophic marker (*HIS3* or *TRP1*) to select for the integration, a regulatable promotor (P_{GAL1} or P_{MET3}), which allowed conditional expression of the target gene depending on the choice of the growth medium and, in addition, a sequence encoding an epitope tag (2 x Protein A or 2 x HA) to allow subsequent immunodetection (section II.3.4.5) and immunoprecipitation (section II.3.5.2) of the produced fusion protein.

II.2.7 Growth curves

To determine the effects of the metabolic depletion of a protein from the cells, cells were grown to mid-logarithmic phase under permissive conditions (galactose-based medium or medium lacking methionine). Then aliquots of the culture were used to inoculate media either providing permissive or repressing conditions (due to the choice of the carbon source or due to the presence or absence of methionine). The media were inoculated to an OD_{600} of 0.05-0.1 and grown at $30^{o}C$. The growth rate was then monitored by measuring the OD_{600} at regular intervals, typically every two hours. Maintenance of logarithmic growth was ensured by diluting the cultures with pre-warmed medium to keep the OD_{600} readings below 0.8.

II.2.8 Yeast two-hybrid screen

The two-hybrid screens in this work were performed following a protocol designed by Fromont-Racine *et al.*, 1997.

II.2.8.1 Bait construction

Two-hybrid bait fusions were constructed in plasmids derived from pAS2ΔΔ or pBTM116 (table II.1.7) using standard recombinant DNA procedures (section II.3.2). All bait constructs were verified by sequencing prior to use. The auto-activation properties of each bait were usually tested either by performing directed two-hybrid mating assays (section II.2.8.8), testing the bait fusion for two-hybrid interaction with a number of prey-fusions available in the laboratory or by performing a small-scale two-hybrid screen. A small scale two-hybrid screen ("mini-screen") was essentially performed as a full scale screen, but only one tenth of the cells were plated out after the mating procedure. From the ratio of cells able to grow on YMM -LWH (cells in which reporter gene expression is supported) to the number of diploids (cells able to grow on YMM -LW) screened, the auto-activating potential of a bait could be evaluated and the number of positive clones expected for a full scale screen could be calculated.

II.2.8.2 The Fromont-Racine yeast libray (FRYL)

The yeast library used in most two-hybrid screens described in this work was constructed by Micheline Fromont-Racine in the laboratory of Pierre Legrain, Institute Pasteur, Paris (Fromont-Racine *et al.*, 1997). Genomic DNA was sonicated and treated with three modification enzymes (mung bean nuclease, T4 DNA polymerase and Klenow-enzyme) to produce blunt ended fragments. Adaptors were ligated to these fragments, producing a 3' overhang and the fragments were then ligated into the pACTIIStop plasmid (table II.1.7), which had been previously cut with *Bam*HI and "filled in" with dGTP by Vent (Exo⁻) polymerase. Library was transformed into *E. coli*

cells MR32, transformed colonies were scraped from the plates, pooled and frozen. The cells were stored at -70°C. The Library DNA was then extracted and transformed into yeast strain Y187. Transformants were scraped from the plates, pooled and aliquoted. Aliquots were stored in 15% glycerol at -70°C.

In this laboratory, the yeast transformation was repeated to increase the stock of two-hybrid library. The yeast library derived from this transformation will be referred to as Edinburgh yeast library (EDYL).

II.2.8.3 Mating and collection of diploid cells

Bait plasmids were transformed (section II.2.4) into their respective carrier yeast strain (CG1945 for pASΔΔ plasmids; L40 for pBTM116 plasmids (table II.1.7) and grown on YMM -W medium (table II.1.2). A bait culture was grown to an OD_{600} of 0.9 to 1.0 units. An aliquot of library cells (1 ml of FRYL or 10 ml of EDYL) was thawed on ice, inoculated into 20 ml of YPDA + tetracycline (tet) (table II.1.2 and section II.1.2.2) and incubated for regeneration at 30°C for 15 min (on shaker, 120 rpm). Bait cells equivalent to 80 OD_{600} units (approximately 8 x 10^8 cells) were mixed with the library plasmid-containing cells and were concentrated onto twelve Millipore filters (45 mm diameter, 0.22 μm). Each filter was washed with 8 ml of fresh YPDA + tet medium and incubated for 4 hours on solid YPDA + tet medium at 30°C. The cells were then washed from the filters with a total of approximately 25 ml YMM -LWH medium (table II.1.2) and collected. The suspension was thoroughly mixed and 50 μl were removed for the controls (These cells were serially diluted to 1:1000 in YMM -LWH medium and 50 μl spread onto each of YMM -L, YMM -W and YMM -WL medium (table II.1.2). The plates were incubated for 2 days at 30°C and the number of colonies on these plates was counted. For the calculation of the mating efficiency and the number of diploid cells screened, see section II.2.8.4). The mated cells were finally spread onto YMM -LWH + tet medium at about 250 μl per plate and incubated at 30°C for 3 days.

II.2.8.4 Calculation of mating efficiency and number of diploids screened

From the number of colonies counted on the control plates (section II.2.8.3), the number of diploid clones (containing both bait and prey plamid) screened as well as the mating efficiency could be calculated:

Diploids screened = colonies on YMM -LW x dilution factor x volume of culture.

In order to ensure complete coverage of the FRYL within one screen, 1.5×10^7 diploids had to be screened.

The mating efficiency was calculated as follows:

mating efficiency (%) = $\dfrac{\text{number of colonies on YMM -LW} \times 100}{\text{number of colonies on YMM -L}}$

II.2.8.5 The X-gal overlay assay

In order to monitor the expression of the β-galactosidase reporter gene, an X-gal overlay assay was performed. Ten millilitres of overlay mix (maintained at $50^{\circ}C$ in a water bath) was pipetted gently onto each plate from the two-hybrid screen and allowed set at room temperature. The colonies were then incubated at $30^{\circ}C$ and examined at regular intervals (usually every 3 hours at early time points, then every 6 hours) for the development of a blue colour.

Overlay mix (final concentrations):
- 0.5% (w/v) agar [a]
- 0.1% (w/v) SDS
- 3.55% (w/v) Na_2HPO_4[a]
- 0.2 % (v/v) ortho-phosphoric acid
- 6% (v/v) dimethyl formamide
- 0.04% (w/v) X-gal (in dimethyl formamide)

[a] autoclaved independently and stored at $65^{\circ}C$ until use.

II.2.8.6 The X-gal filter-lift assay

Filter-lift assays were performed as described in Transy and Legrain (1995). Filter-lift assay solution for detection of β-galactosidase activity was prepared fresh from stock solutions immediately prior to use. The cells were transferred to Hybond-C extra filters (Amersham), and the filters immersed in liquid nitrogen for 5 seconds. Filters were placed (cell-side up) onto Whatman 3MM paper soaked in assay solution, incubated at $30^{o}C$ and examined at regular intervals (usually every hour). The reaction was stopped by transferring the filters onto Whatman 3MM paper soaked in 1 M Na_2CO_3 for 1 minute. The filters were then transferred onto Whatman paper soaked in sterile water for 1 min and air dried.

Z-buffer: 100 mM $NaPO_4$, pH 7.5

                     10 mM KCl

                     1 mM $MgSO_4$ x 7 H_2O

                     sterilised by autoclaving

Assay solution: 0.27% (v/v) β-mercaptoethanol

                     0.04% (w/v) X-gal (in dimethyl formamide)

                     Prepared in Z-buffer

II.2.8.7 Analysis of positive colonies from two-hybrid screens

Colonies which were identified as "positive" with respect to the expression of the reporter gene(s) were analysed to identify the library-plasmid responsible for the two-hybrid interaction. Plasmid DNA was rescued from the yeast cells by the method described in section II.3.2.4 and transformed (section II.2.4.2) into electro-competent MC1066 cells (table 2.1.6). These cells were plated onto M9 -Leu medium (table II.1.1), which allowed growth of only those cells carrying a library plasmid. The plasmid DNA was propagated in these cells and prepared as described in section II.3.2.2. The insert size was determined by restriction digest of the plasmid with

*Bam*HI (which cuts in the library adaptor sequence at both ends of the insert) and gel electrophoresis (section II.3.2.10). The fusion between the vector sequence and the insert was determined by sequencing (section II.3.2.16). Finally, the identity of the genomic insert was revealed using the *Saccharomyces* Genome Database (SGD; section II.4).

Alternatively to the above described plasmid rescue strategy, in some two-hybrid screens the identity of the library plasmid insert was determined via PCR-amplification of the insert directly from the yeast colonies (section II.3.2.8.2), subsequent gel-purification of the PCR product using the QIAquick Gel Extraction Kit (section II.3.2.11) and sequencing (section II.3.2.16) of the fragment.

II.2.8.8 Directed two-hybrid mating assay

A directed mating strategy was used to test for potential two-hybrid interactions between cloned bait(s) and prey fusions. Yeast strains CG1945 or L40 (table II.1.5) transformed with bait plamids and yeast strain Y187 transformed with prey plasmids were grown on selective media (table II.1.2). Bait and prey strains were then mated by replica-plating onto YPDA and incubation overnight at 30°C. The resulting diploids were grown on medium selecting for both bait and prey plasmids. The diploid colonies were then suspended in 100 μl of sterile distilled water filled into wells of a microtiter plate. The cells were then transferred onto medium selecting for a successful two-hybrid interaction using a pronged metal inoculator. The stringency of the interaction was examined by assaying growth of the diploids on selective medium containing different concentrations of 3AT (usually 0-50 mM).

II.3 Molecular biology methods

II.3.1 General methods

II.3.1.1 Spectrophotometric determination of nucleic acid concentrations

The concentration of DNA or RNA was determined by measuring the absorption of diluted DNA solutions at 260 nm using a Cecil CE 2040 spectrophotometer and a quartz cuvette. For double stranded DNA an OD_{260} value of 1.0 represents a DNA concentration of 50 µg/ml. For single stranded RNA an OD_{260} value of 1.0 represents an RNA concentration of 40 µg/ml. DNA purity was determined by measuring the absorption at wavelengths of 260 and 280 nanometres. Protein free preparations of DNA or RNA should give OD_{260}:OD_{280} ratios of 1.8 or 2.0, respectively.

II.3.1.2 Extraction with phenol:chloroform:isoamylalcohol

Nucleic acids were separated from protein in preparations by adding an equal volume of phenol:chloroform:isoamylalcohol (P/C/I) (25:24:1), vortexing 10 seconds and centrifuging at 14,000 rpm for 5 minutes. The upper, nucleic acid containing, aqueous phase was transferred to a fresh tube.

II.3.1.3 Precipitation of nucleic acids

Nucleic acids were precipitated from solutions by addition of 1/10 volume of 3 M NaOAc (pH 5.2) and 2.5 volumes of ethanol (abs.) and freezing for 20 minutes at -70°C. The nucleic acids were sedimented by centrifugation at 14,000 rpm for 20 minutes at 4°C with the pellet washed in 70% (v/v) ethanol. The pellet was washed again in ethanol (abs.) and air dried.

II.3.2 DNA methods

II.3.2.1 Small scale preparation of plasmid DNA from *E. coli*

This method is based on the alkaline lysis method of Zhou *et al*. (1990). 3 ml of LB medium (table II.1.1) supplemented with the appropriate antibiotic (section II.1.2.2) was inoculated with a single colony of the plasmid-bearing *E. coli* strain and incubated overnight at 37°C with constant shaking. One and a half millilitres of the culture were centrifuged at 14,000 rpm for 30 seconds, the supernate decanted and the cell pellet resuspended in 150 μl of TENS solution by vortexing. Then, 75 μl of 3 M NaOAc (pH 5.2) was added and the sample vortexed again. Cell debris and chromosomal DNA were pelleted by centrifugation at 14,000 rpm for 5 minutes and subsequently removed with a sterile toothpick. The plasmid DNA was then extracted with P/C/I (section II.3.1.2), precipitated, washed and dried (section II.3.1.3). The cell pellet was resuspended in 25 μl of sterile destilled water.

TENS solution: 10 mM Tris-HCl, pH 7.5

1 mM EDTA

100 mM NaOH

0.5% (w/v) SDS

II.3.2.2 Small scale preparation of plasmid DNA by spin column

For automated DNA sequencing (section II.3.2.16) plasmid DNA was prepared using the QIAprep kit (Quiagen), following the manufacturers' guidelines. DNA was isolated from 4.5 ml of *E. coli* cell culture, eluted with 100 μl sterile distilled water and stored at -20°C.

II.3.2.3 Large scale preparation of plasmid DNA from *E. coli*

One hundred mililitres of LB medium (table II.1.1) supplemented with the appropriate antibiotic (section II.1.2.2) was inoculated with a single colony of a plasmid-bearing strain and incubated overnight at 37°C with constant shaking. Cells were harvested by centrifugation at 6000 x g for 10 min (4°C, Beckman JLA10.500 rotor) and resuspended in 4 ml of GTE. Eight millilitres of lysis solution was added, mixed gently and incubated on ice for 10 minutes. Six millilitres of ice-cold KOAc (3 M KOAc, 11.5% (v/v) acetic acid) was added. This cell suspension was mixed gently and incubated on ice for 5 minutes. Cell debris and chromosomal DNA were pelleted by centrifugation at 20,000 x g for 10 minutes (4°C, Beckman JA25.50 rotor) in a polypropylene tube and the supernate decanted to a fresh 30 ml corex tube. Eleven millilitres of isopropanol was added, mixed and the tube centrifuged at 10,000 rpm for 10 minutes at room temperature (Beckman JA25.50 rotor). The pellet was washed with 70% (v/v) ethanol and allowed to dry. The dried pellet was fully resuspended in 1 ml 1 x TE (table II.1.3). and 1.13 g of CsCl and 75 µl of ethidium bromide (10 mg/ml stock solution) added and mixed. The mixture was transferred to a Quick-seal tube (11 x 32 mm, Beckman) and centrifuged at 90,000 rpm for 16 hours (18°C, Beckman TLA120.2 rotor). The plasmid DNA band was collected from the CsCl gradient with a syringe and needle and residual ethidium bromide removed by three CsCl-saturated isopropanol extractions. Plasmid DNA was precipitated by addition of 4 volumes of ethanol (abs.) and 1 volume of 1 M NH_4OAc and centrifugation at 14,000 rpm for 10 minutes at room temperature. The DNA pellet was washed with 70% (v/v) ethanol and dried under vacuum. The dried pellet was resuspended in 200 µl of 1 x TE. A second precipitation was performed by addition of 2 volumes of ethanol (abs.) and 0.1 volumes of 3 M NaOAc (pH 5.2) and centrifugation at 14,000 rpm for 10 minutes at room temperature. The DNA pellet was washed with 70% (v/v) ethanol, dried under vacuum and resuspended in 500 µl of sterile destilled water. The plasmid DNA was stored at -20°C.

Lysis solution: 0.2 N NaOH

                                 1% (w/v) SDS

GTE: 50 mM Glucose

                        25 mM Tris-HCl

                        10 mM EDTA

II.3.2.4 Plasmid rescue from yeast cells

Two mililitres of the appropriate selective medium (table II.1.2) were inoculated with a single yeast colony and incubated on a rotating wheel overnight at 30°C. One and a half millilitre of the cultures were transferred to a microcentrifuge tube and the cells were sedimented by centrifugation at 14,000 rpm for 5 minutes. Two hundred microlitres of extraction buffer and 400 μl of glass beads were added. Tubes were spun briefly with open lids (to remove residual glass beads from the edge of the tube, which do prevent closing the tube properly). Two hundred microlitres of P/C/I was added and the cells were vortexed for 7 minutes. The sample was then centrifuged for 5 minutes at 15,000 rpm. One hundred and sixty microlitres of the supernate was transferred to a fresh tube and 600 μl of ethanol (abs.)/7.5 M NH_4OAc (6:1) was added and the sample vortexed for 10 seconds. The DNA was pelleted by centrifugation for 5 minutes (15,000 rpm, 4 $^{\circ}$C), washed with 70% ethanol and ethanol (abs.) and air dried. The pellet was resuspended in 10 μl of sterile distilled water.

Extraction buffer: 2% (v/v) Triton-X-100

                        1% (w/v) SDS

                        100 mM NaCl

                        10 mM Tris-HCl, pH 8.0

                        1 mM EDTA, pH 8.0

II.3.2.5 Preparation of yeast genomic DNA

Yeast cells were grown overnight in 5 ml of the appropriate medium (table II.1.2) and transferred to microcentrifuge tubes. The cells were sedimented by centrifugation at 14,000 rpm for 15 seconds, resuspended in 0.5 ml of buffer A and incubated at $37^{o}C$ for 1 hour. The spheroblasts were centrifuged for 1 minute at 14,000 rpm and then resuspended in 0.5 ml of buffer B. Fifty microlitres of 10% (w/v) SDS was added, the sample mixed well and placed in a water bath at $65^{o}C$ for 30 minutes. Then 200 μl of 5 M KOAc was added and the sample was incubated on ice for 60 minutes. The tube was centrifuged at 14,000 rpm for 5 minutes and the supernate transferred to a fresh tube. One volume of isopropanol was added, the sample mixed and incubated for 5 min at room temperature. The DNA was pelleted by centrifugation at 14,000 rpm for 30 seconds and air dried. The pellet was resuspended in 300 μl of 1 x TE (table II.1.3) and the DNA was precipitated again, this time by addition of 1/10 volume of 3 M NaOAc (pH 5.2) and 200 μl of isopropanol and centrifugation at 14,000 rpm for 30 seconds. The pellet was air dried and resuspended in 100-300 μl 1 x TE.

Buffer A: 0.9 M Sorbitol
              0.1 M EDTA
              50 mM DTT
              500 U lyticase/ml

Buffer B: 50 mM Tris-HCl, pH 7.5
              50 mM EDTA

II.3.2.6 Restriction digest of DNA

Restriction endonuclease digestion of DNA was typically performed in volumes of 10 to 100 μl. These contained the requisite quantity of DNA and the appropriate buffer (as supplied by the manufacturer) at 1 x concentration. Between 2 and 5 units of restriction

enzyme were added, with the restriction enzyme volume kept below 10% of the total volume. The digestion was incubated at the temperature recommended by the supplier, typically for a period of 2-4 hours. The products of the digestion were either analysed by agarose gel electrophoresis (section II.3.2.10) or extracted with P/C/I (section II.3.1.2) and ethanol-precipitated (section II.3.1.3) for further manipulations.

II.3.2.7 Removal of phosphates from DNA ends

Plasmid DNA digested with restriction endonucleases (section II.3.2.6) was incubated with 5 units of calf intestinal (alkaline) phosphatase (CIP) (for 20 μl reactions) for 30 minutes at $37^{o}C$ to remove terminal phosphate groups and thereby prevent the recircularisation of the vector DNA.

II.3.2.8 Amplification of DNA using the polymerase chain reaction

II.3.2.8.1 DNA amplification from prepared plasmid or genomic DNA

Specific regions of DNA were amplified using the polymerase chain reaction (PCR). Template DNA was either a small quantity of plasmid DNA (typically 10 ng in 1 μl) or one microlitre of genomic DNA, prepared as desribed in section II.3.2.5. Oligonucleotides used for priming the reaction are listed in table II.1.6. A typical 100 μl reaction using genomic DNA as template was set up as follows:

10 x polymerase buffer	10 μl
100 mM $MgCl_2$*	2 μl
2.5 mM dNTPs (dATP, dCTP, dGTP, dTTP)	10 μl
Oligonucleotide primer 1 (10 pmol/μl)	10 μl
Oligonucleotide primer 2 (10 pmol/μl)	10 μl
Template DNA (section II.3.2.5)	1 μl
DNA polymerase (2 U/μl)	1 μl

Sterile distilled water	56 µl

* $MgCl_2$ concentration was titrated to 2 and 4 mM as required.

(when plasmid DNA was used as template, normally 1 µl (10ng/µl) was used)

The reaction mix was overlaid with a drop of mineral oil (Sigma) to prevent evaporation during the reaction cycles. All PCRs were carried out in a Hybaid Thermal Reactor or in a PTC-100 Hot Lid reactor from Genetic Research Instrumentation Ltd (in this case the mineral oil was avoided). The reactor was programmed depending on the length of the desired product and the annealing temperature of the oligonucleotide primers used. A typical programme is described below.

Step 0. Denaturation:	95°C	3 minutes
Step 1. Denaturation:	95°C	30 seconds
Step 2. Annealing:	40-60°C*	30 seconds
Step 3. Extension:	72°C	60-180 seconds
30 cycles of Step 1- Step 3.		
Step 4: Final extension:	72°C	3 minutes

The annealing temperature was calculated for oligonucleotides of up to 20 bases in length with the formula: annealing temperature = 4 x ((G+C) + 2 x (A + T)) - 5°C.

II.3.2.8.2 DNA amplification directly from yeast colonies

For amplification of a particular DNA region directly from yeast, one large single colony for each PCR was resuspended in 0.02 N NaOH, boiled for 5 minutes and put on ice. Two microlitres of this suspensions were transferred to fresh 0.5 ml microcentrifuge tubes on ice. A premix for 14 PCRs was prepared as follows:

Premix: 344.4 μl sterile distilled water

44.8 μl 10 x PCR buffer

9.8 μl 10 mM dXTP

7 μl oligonucleotide primer 1 (15-20 pmol/μl)

7 μl oligonucleotide primer 2 (15-20 pmol/μl)

The premix was mixed well and kept on ice until use.

The thermal reactor was pre-heated to $94^{o}C$, 7 μl Taq polymerase (5 U/μl, Boehringer Mannheim) was added to the premix and 30 μl of premix were distributed to each 2 μl of cell suspension. The tubes were then transferred to a PTC-100 Hot Lid reactor (Genetic Research Instrumentation Ltd) programmed as follows:

Step 0: $94^{o}C$ 3 minutes

Step 1: $94^{o}C$ 30 seconds

Step 2: $55^{o}C$ 90 seconds

Step 3: $72^{o}C$ 3 minutes

31 cycles of step 1- step 3.

Step 4: $72^{o}C$ 5 minutes

3 μl of the PCR was then analysed by agarose gel electrophoresis (section II.3.2.10). The PCR product was purified as described in section II.3.2.9.

II.3.2.9 Purification of PCR products

If the performance of a polymerase chain reaction led to the production of a single product, the product was purified from oligonucleotide primers, nucleotides, polymerase and salts using a QIAquick PCR purification column (Quiagen) as recommended by the manufacturer. The DNA was typically eluted from the column with 30 μl of sterile distilled water.

If the PCR produced additional products apart from the desired one, the total PCR (section II.3.2.8) was subjected to agarose gel electrophoresis (section II.3.10) and the desired band isolated and purified from the gel slice as described in section II.3.2.11.

II.3.2.10 Agarose gel electrophoresis

DNA fragments produced by restriction endonuclease digest or generated by amplification in a PCR were analysed in 0.7-2.0% (w/v) agarose gels. Gels were prepared by melting agarose in 1 x TAE buffer (table II.1.3) and adding ethidium bromide to a final concentration of 0.5 μg/ml. Samples to be analysed were mixed with 1/6 volume of loading buffer prior to loading.

6 x loading buffer: 0.25% bromophenol blue

0.25% xylene xyanol FF

15% Ficoll (Type 400) in water

stored at room temerature

II.3.2.11 Isolation of DNA from agarose gels

To isolate and purify specific DNA bands from agarose gels, the QIAquick Gel Extraction Kit (Qiagen) was used, as recommended by the manufacturers' protocols. DNA fragments were separated by agarose gel electrophoresis (section II.3.2.10) and the bands visualised on a UV transilluminator. The band to be purified was excised with a clean razor blade and purified. DNA was typically eluted in 30 μl of sterile distilled water and stored at -20°C.

II.3.2.12 Creation of blunt ended DNA fragments

II.3.2.12.1 Filling in recessed 3' termini

To fill in recessed 3' termini of DNA fragments created in restriction digests, typically 5 µl of 2.5 mM dATP, dCTP, dGTP and dTTP as well as 1 µl of Klenow fragment of DNA polymerase I (5 U/µl) was added to 40 µl of restriction digest sample (in which the restriction enzyme had previously been heat-inactivated). If the DNA to be filled in was suspended in sterile distilled water, the reaction had to be supplemented with the appropriate 10 x Klenow DNA polymerase reaction buffer supplied by the manufacturer. The samples were incubated at $37^{o}C$ for 30 min and the Klenow enzyme was heat-inactivated by incubation at $75^{o}C$ for 20 minutes.

II.3.2.12.2 Removal of 3' overhangs

Three prime overhangs produced by restriction digests were blunted if neccessary by adding 1 µl of T4 DNA polymerase (3 U/µl) to the previously heat-inactivated samples and incubation for 20 minutes at $37^{o}C$. The 3' to 5' exonuclease activity of the T4 DNA polymerase cuts off the 3' overhangs and blunt ends are created. The enzyme was heat-inactivated by incubation at $75^{o}C$ for 20 minutes. The DNA was precipitated (section II.3.1.3) prior to further manipulation.

II.3.2.13 Ligation of DNA molecules

Ligations were typically performed in a final volume of 15 µl, containing 0.5-1.0 µg of total DNA, 1 x ligation buffer, 1 mM ATP and 0.5 units of Fast-Link DNA ligase (Epicentre Technologies). Vector and insert DNA were present in approximately 1:3 to 1:5 ratio. Reactions were allowed to proceed at room temperature for 20 minutes and stopped by heat-inactivation of the enzyme for 15 min at $70^{o}C$. The reaction mix was ethanol precipitated and the pellet washed (section II.3.1.3). The pellet was

resuspended in 10 µl sterile distilled water. Typically, 1 µl was used in the subsequent *E. coli* transformation (section II.2.4.2).

II.3.2.14 Radio-labelling of DNA fragments by random priming

DNA fragments were radio-labelled using the random priming method of Feinberg and Vogelstein (1983; 1984). The DNA fragment to be labelled was prepared by restriction digest (section II.3.2.6) or generated by PCR (section II.3.2.8) and gel-purified (section II.3.2.11). Sterile destilled water was used to increase the volume of the DNA solution (containing 50-100 ng of DNA) to 27 µl, and the solution was then boiled for 5 min and put on ice. The random priming reaction was prepared as follows and incubated for 6-8 hours at room temperature:

Denatured DNA in water	27 µl
OLB*	10 µl
BSA (2 mg/ml)	10 µl
[α-^{32}P] dCTP (~5000 Ci/mmol)	3 µl
Klenow DNA polymerase I (5 U/µl)	0.5 µl

* OLB consists of a 1:2.5:1.5 mixture of solutions A:B:C prepared as follows and stored at -20$^\text{O}$.

Solution A: 18 µl β–mercaptoethanol plus 5 µl each of 100 mM dATP, dGTP and dTTP diluted to 1 ml with solution O. Stored at -20$^\text{O}$.

Solution B: 2 M HEPES buffer, pH 6.6 (adjusted with NaOH). Stored at -20$^\text{O}$.

Solution C: Random hexadeoxyribonucleotides (Pharmacia) at 90 OD units/ml in sterile distilled water. Stored at -20$^\text{O}$.

Solution O: 1.25 M Tris-Cl, pH 8.0, 0.125 M MgCl$_2$

Random primed DNA was separated from unincorporated nucleotides using a NAP-5 column (Pharmacia) according to the manufacturers' protocol.

II.3.2.15 End-labelling of oligonucleotides

Oligonucleotides were labelled in a 20 μl reaction at 37°C for 45 minutes using T4 polynucleotide kinase (PNK, New England Biolabs). The reaction mix was prepared as follows:

Oligonucleotide	10 pmol (volumes vary)
10 x PNK buffer	2 μl
T4 PNK (10 U/μl)	0.5 μl
$[\gamma\text{-}^{32}P]ATP$ (~5000 Ci/mmol)*	2 μl

adjusted to 20 μl with sterile, distilled water

II.3.2.16 DNA sequencing

Plasmid DNA to be sequenced was prepared using QIAprep spin columns (section II.3.2.2) and quantitated by visualisation on in an agarose gel (II.3.2.10). Reactions were performed with the dRhodamine terminator cycle sequencing kit (Perkin Elmer) in a PTC-100 Hot Lid reactor (Genetic Research Instrumentation Ltd). A typical reaction mix was set up as follows:

Template DNA (~400 ng)	4 μl
Terminator mix	4 μl
Primer (3.2 pmol)	2 μl

Twenty five cycles as described below were performed:

Step 1: 96°C for 30 seconds

Step 2: 50°C for 15 seconds

Step 3: 60°C for 4 minutes

The DNA was precipitated from the reaction mix by adding 50 μl of ethanol (abs.) and 2 μl of 3 M NaOAc, pH 5.2, incubating on ice for 15 minutes followed by centrifugation at 14,000 rpm for 20 minutes at 4°C. The pellet was washed with 250 μl of 70% (v/v) ethanol, washed again with ethanol (abs.) and air dried. The samples were run by Nicola Preston (University of Edinburgh) on an ABI PRISM 377 DNA sequencer and the sequence analysed using the Gene Jockey II programme on a Macintosh computer.

II.3.2.17 Southern blot analysis

Yeast genomic DNA was prepared (section II.3.2.5) digested with the required restriction enzyme (section II.3.2.6) and resolved in an agarose gel (section II.3.2.10). The gel was immersed in denaturing buffer (0.5 M NaOH/1.5 M NaCl) for 45 minutes with gentle agitation, then transferred to neutralisation buffer (1.5 M NaCl, 0.5 M Tris-HCl (pH 7.5), 1 mM EDTA) for 45 minutes with gentle agitation.

The DNA was then transferred from the gel to a nylon membrane as follows:
Hybond-N nylon membrane (Amersham) and four sheets of Whatman 3MM paper were cut to the same size as the gel. One strip of Whatman paper was saturated with 20 x SSC (table II.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 20 x SSC. The pre-treated gel was placed on top of the saturated Whatman paper and Saran wrap was carefully placed on top of regions not covered by the gel to prevent unneccessary evaporation. The nylon membrane was placed on top of the gel, and the four sheets of pre-cut Whatman 3MM paper were saturated in 20 x SSC and placed on top of the membrane. On top of this arrangement, 4 cm of dry paper towels were placed and the whole structure weighted to provide even pressure. Transfer was allowed to take place overnight. After the transfer was complete, the membrane was briefly rinsed in 20 x SSC, blotted dry, and UV-irradiated in a Stratagene UV Stratalinker using the autocrosslink option (120 mjoules, 254 nm) to immobilise the DNA to the filter.

Hybridisation of randomly-labelled probes to the nylon membrane was then performed in adaptation to the method described by Church and Gilbert (1984). The hybridisations were performed in Hybaid "Hybridiser" ovens. The nylon membrane was pre-hybridised in 20 ml of SES1 buffer for 2 hours at 65°C to prevent non-specific hybridisation of the probe to the membrane. Fresh SES1 buffer was added to the membrane immediately before the addition of the radiolabelled probe. The labelled probe (section II.3.2.14) was added to the SES1 buffer and the incubation continued overnight at 65°. The probe was decanted off and stored at -20°C for possible re-use. The membrane was then washed with SES2 buffer for 20 min at 60°C four times, with fresh pre-warmed buffer for each wash. The membrane was blotted dry, placed between Saran wrap, and the result of the experiment was visualised by autoradiography.

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

SES2 buffer:	5% (w/v) SDS
	1 mM EDTA
	40 mM Sodium phosphate buffer, pH 7.2

For removal of the probes from the membrane (which allowed reprobing), 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.

II.3.3 RNA methods

II.3.3.1 Total RNA preparation from yeast

Total yeast RNA was prepared using the method of Schmitt *et al.* (1990). Ten millilitres of the appropriate medium (table II.1.2) were inoculated with a single yeast colony and incubated at $30^{o}C$. The cells were harvested by centrifugation at 3,500 rpm for 3 minutes (Mistral 1000 centrifuge) and resuspended in 400 µl of EA buffer (50 mM NaOAc, 10 mM EDTA). Forty microlitres of 10% SDS were added to the cell suspension and vortexed. Five hundred microlitres of phenol, equilibrated with EA buffer, was added, the sample mixed and incubated at $65^{o}C$ for 4 minutes. The samples were snap frozen on dry ice and allowed to thaw prior to centrifugation at 14,000 rpm for 2 minutes. The aqueous layer was removed and extracted as described in section II.3.1.2. The aqueous layer was again removed and precipitated as described in section II.3.1.3. The precipitated RNA was dried under vacuum and resuspended in 30 µl of sterile distilled water.

II.3.3.2 Northern blot analysis of mRNA

Denaturing gel electrophoresis: Gels (1% (w/v)) were prepared by dissolving 1.5 g of agarose in 110 ml of water, allowing to cool slightly and adding 15 ml of 10 x HEPES buffer. The mixture was cooled to $50^{o}C$ and 25 ml of 37% (v/v) formaldehyde added together with 10 µl of 10 mg/ml ethidium bromide. Total RNA was prepared as described in section II.3.3.1, mixed with 3 x loading buffer and heated to $65^{o}C$ for 5 minutes immediately prior to loading. Electrophoresis was performed in 1 x HEPES running buffer at 5 V/cm, with the buffer circulated slowly by a peristaltic pump.

> 10 x HEPES buffer: 0.5 M HEPES (adjusted to pH 7.8 with KOH)
> 10 mM EDTA

3 x loading buffer: 50% (v/v) Formamide

                                  6% (v/v) Formaldehyde

                                  0.4 x HEPES buffer

                                  0.075% (w/v) Bromophenol blue

                                  0.075% (w/v) Xylene cyanol

                                  10% (v/v) Glycerol

Alternatively, the electrophoresis was performed as described in section II.3.2.10 within an 1.2 % (w/v) agarose gel. Denaturing conditions were yielded by mixing the 10 μl (10 ng of RNA) samples with 5 μl loading buffer and heating the samples for 10 minutes at 65°C prior to loading. The gel was run for 3.5 hours at 80 V.

Loading buffer: 16 ml Formamide

                              4 ml 50 mM Tris-HCl, pH 7.6

                              0.05 g Xylene xyanol

                              0.05 g Bromophenol blue

The RNA was then transferred electrophoretically from the gel to a Hybond-N nylon membrane at 60 V for 60 minutes in 0.5 x TBE (table II.1.3). After the transfer the RNA was cross-linked to the membrane as described in section II.3.2.17.

The hybridisation of randomly labelled DNA probes (section II.3.2.14) to the RNA were performed as described in section II.3.2.17.

Oligonucleotides, which had been end-labelled as described in section II.3.2.15, were hybridised as follows: the nylon membrane was pre-hybridised in 20 ml of formamide hybridisation buffer for 2 hours at 37°C. Ten millilitres of fresh formamide hybridisation buffer was added to the membrane immediately before the addition of the end-labelled probe. The labelled probe was added to the membrane and the incubation continued overnight at 37°C. The membrane was washed with 2 x SSC (table II.1.3) for 5 minutes at 37°C four times, with fresh buffer for each wash. The membrane was not allowed to dry and was placed in between Saran wrap, and the result was visualised by autoradiography.

Formamide hybridisation buffer:	50% (v/v) Formamide
	5 x SSPE
	5 x Denhardts solution
	1% (w/v) SDS
	200 μg/ml Salmon sperm DNA
20 x SSPE:	3.6 M NaCl
	0.2 M NaH_2PO_4
	0.02 M EDTA
	pH adjusted to 7.4 with NaOH
100 x Denhardts solution:	2% (w/v) BSA
	2% (w/v) Ficoll 400
	2% (w/v) Polyvinylpyrrolidone (PVP)

II.3.3.3 Primer extension assay

An oligonucleotide was labelled by end-labelling. A reaction mix was set up as detailed below and the reaction incubated at 37°C for 30 minutes:

Oligonucleotide (10 pmol/μl)	2 μl
10 x kinase buffer	1.5 μl
100 mM DTT	1.5 μl
[γ-^{32}P] ATP	2.5 μl
T4 PNK (10 U/μl)	1 μl
Sterile distilled water	6.5 μl

The labelled oligonucleotide was precipitated by adding 90 μl of ethanol (abs.), 8 μl of 2 mg/ml glycogen and 11.5 μl of 7.5 M NH_4OAc and kept at -70°C overnight. The oligonucleotide was then pelleted by centrifugation at 14,000 rpm for 30 minutes at

4°C. The DNA pellet was washed with 70% (v/v) ethanol, allowed to dry and resuspended in 50 μl of sterile distilled water.

Primer annealing - Yeast total RNA was prepared as described in section II.3.3.1. Typically, two labelled oligonucleotides were annealed to the target RNA, one to assay the levels of the RNA species under investigation, the other one to act as control for RNA loading. Ten micrograms of RNA was mixed with 2 μl of 5 x SS hybridisation buffer and 1 μl of each of the labelled oligonucleotides in a total volume of 12 μl. The reaction was placed at 80°C for 5 minutes, before being incubated at 46°C for 90 minutes.

Primer extension - Forty microlitres of 1.25 x RT buffer (pre-warmed at 46°C) containing 0.5 μl of reverse transcriptase (20 U/μl) and 0.5 μl of RNasin was added to the annealed RNA sample and the reaction was incubated at 46°C for 40 minutes. After this time, 6 μl of 1 M NaOH and 1 μl of 0.5 M EDTA was added and the reaction incubated at 55°C for a further 45 minutes.

Polyacrylamide gel electrophoresis - The DNA molecules produced by the reverse transcriptase were precipitated by the addition of 250 μl of ethanol (abs.), 2 μl of 2 mg/ml glycogen, 30 μl of 7.5 M NH_4OAc and 6 μl of 1 M HCl, followed by incubation for 30 min at -70°C and centrifugation at 14,000 rpm for 15 minutes at 4°C. The DNA pellet was allowed to dry and was then resuspended in formamide loading buffer. The sample was heated to 80°C for 3 minutes immediately before being loaded onto a 6% (w/v) polyacrylamide gel (Sequagel-6:Sequagel complete buffer reagent 4:1).

The gel was run at 24 W in 1 x TBE (table II.1.3) for approximately 40 minutes. The gel was covered with Saran wrap and exposed to autoradiography film.

10 x kinase buffer: 700 mM Tris-HCl

                                          100 mM $MgCl_2$

5 x hybridisation buffer:	1.5 M NaCl
	50 mM Tris-HCl, pH 7.5
	10 mM EDTA

1.25 x RT buffer:	12.5 mM Tris-HCl
	12.5 mM DTT
	7.5 mM $MgCl_2$
	1.25 mM each of dATP, dCTP, dGTP, dTTP

formamide loading buffer:	51% (v/v) Formamide
	20 mM EDTA
	0.3% (w/v) Xylene xyanol
	0.3% (w/v) Bromophenol blue

II.3.3.4 *In vitro* transcription

Preparation of DNA template - RNase free plasmid DNA p283 (table II.1.7) was linearised by digestion (section II.3.2.6) with *Bam*HI. Approximately 1 μg of DNA was digested in a volume of 20 μl. Following digestion, 1 μl of the template DNA was used directly in an *in vitro* transcription reaction.

Transcription reaction - the transcription reaction detailed below was set up in a microcentrifuge tube and was incubated at 37°C for 30 minutes. The transcript could be stored for up to 5 days at -20°C.

Linearised template DNA	1 μl
T7 buffer	1.5 μl
10 mM each of ATP, CTP, GTP	1 μl
400 μM UTP	1 μl
[α-^{32}P]UTP	1 μl
Sterile, distilled water	11.5 μl
RNasin	0.3 μl
T7 RNA polymerase (61U/μl)	0.5 μl

78

T7 buffer:	0.4 M Tris-HCl, pH 8.0
	0.1 M $MgCl_2$
	0.1 M DTT
	0.1 M NaCl

Precipitation of transcript - the produced actin transcript was diluted and used directly in an *in vitro* splicing reaction. Alternatively, the transcript was precipitated prior to use: the volume of the reaction was increased to 100 µl by addition of sterile, distilled water. The diluted transcript was then P/C/I extracted (section II.3.1.2) and precipitated with ethanol (section II.3.1.3). The pellet was dried under vacuum and resuspended in 50 µl of sterile distilled water.

II.3.4 Protein methods

II.3.4.1 Crude extraction of total cellular protein from yeast

Five millilitres of the appropriate medium (table II.1.2) was inoculated with a single yeast colony and incubated at $30^{\circ}C$ overnight. Twenty five millilitres of fresh medium was inoculated to an OD_{600} of 0.1 and incubated again. When the cells had reached an OD_{600} of 0.5-0.6, the cells were harvested by centrifugation at 3,500 rpm (Mistral 1000 centrifuge) for 5 minutes at $4^{\circ}C$. The pelleted cells were washed with 25 ml of ice-cold, sterile, distilled water and snap-frozen on dry ice. The cell pellet was thawed by addition of 100 µl of pre-warmed ($60^{\circ}C$) cracking buffer per 7.5 OD_{600} units. The cracking buffer was supplemented with 1% (v/v) β-mercaptoethanol and 1 x LPC protease inhibitor cocktail, both added immediately prior to use. The cell suspension was transferred to an microcentrifuge tube containing 100 µl of glass beads (425-600 µm) per 7.5 OD_{600} units and incubated at $70^{\circ}C$ for 10 minutes. The suspension was then vortexed for 1 minute and subjected to centrifugation at 14,000 rpm for 5 minutes at $4^{\circ}C$. The supernate was removed to a fresh tube and the extract stored at $-70^{\circ}C$.

Cracking buffer: 8 M Urea

5% (w/v) SDS

40 mM Tris-HCl, pH 6.8

0.1 mM EDTA

0.4 mg/ml Bromophenol blue

1000 x LPC protease inhibitor cocktail:

10 mg/ml	Leupeptin
10 mg/ml	Pepstatin
10 mg/ml	Chymostatin

prepared in DMSO

II.3.4.2 Large scale extraction of total cell protein (splicing extract)

Total yeast cell protein preparations (splicing extracts) were made using the method of Lin *et al.* (1985). Two litres of the appropriate medium (table II.1.2) were inoculated with cells of a mid-logarithmic phase culture and the cells grown overnight with constant shaking to an OD_{600} of 0.5-1.5. The cells were harvested by centrifugation at 5000 rpm for 5 minutes ($4^{o}C$; Beckman JLA 10.500 rotor), resuspended in 50 ml of 50 mM KPO_4 and transferred to a falcon tube. The cells were pelleted again by centrifugation at 3,500 rpm for 3 minutes at room temperature (Mistral 1000) and were resuspended in 40 ml of lyticase buffer. Two thousand units of Lyticase were suspended in 1 ml of lyticase buffer and added to the cell suspension, which was incubated (in a 500 ml flask) under constant shaking (80 rpm) at $30^{o}C$ for normally 30-60 minutes. When approximately 70-80% of the cells have formed spheroplasts, as determined microscopically, the cells were harvested by centrifugation at 3,000 rpm for 5 minutes at room temperature (Mistral 1000 centrifuge). The supernate was discarded and the spheroplasts were washed twice with 30 ml of 1.2 M sorbitol and once in 30 ml of ice-cold SB-3 buffer. After each wash the spheroplasts were harvested by centrifugation at 3,000 rpm for 5 minutes at room temperature (Mistral

1000 centrifuge) and resuspended gently by using a sterile, sealed Pasteur pipette. After the wash with SB-3 buffer, the pellet was weighted and resuspended in approximately 1 ml of buffer A for each gram of cell pellet. The suspension was then transferred to a pre-chilled Dounce homogeniser and was homogenised by 13 slow, rotation strokes of the tight fitting pestle, with the homogeniser maintained on ice. The lysate was then transferred to a chilled sterile beaker and 1/9 volume of 2 M KCl was added dropwise, with continuous gently mixing of the lysate for 30 minutes at $4^{o}C$. The lysate was then transferred to a chilled polycarbonate centrifuge tube and centrifuged at 17,000 rpm for 30 minutes at $4^{o}C$ (Beckman JA 25.50 rotor, pre-chilled to 4°). Without disturbing the lipid layer, the supernate was transferred to a chilled Ultra plus polycarbonate tube (Nalgene) and centrifuged at 40,000 rpm for 70 minutes at $4^{o}C$ (Beckman 70.1Ti rotor (pre-cooled to $4^{o}C$)). The final supernate was centrifuged at 14,000 rpm for 10 min at $4^{o}C$ and the supernate was then transferred to a Slide-A-Lyzer dialysis cassette (10 kDa cut-off, Pierce) and dialysed against two changes of 1.5 l of ice-cold buffer D at $4^{o}C$ over a period of 4 hours. The extract was then transferred to pre-chilled microfuge tubes centrifuged for 10 minutes at 14,000 rpm ($4^{o}C$) and aliquoted into fresh pre-chilled microfuge tubes and stored at -70°. The extracts were thawed slowly on ice prior to use.

Lyticase buffer:	1.2 M Sorbitol
	50 mM KPO_4, pH 7.5
	30 mM DTT (added immediately before use)
SB-3 buffer:	1.2 M Sorbitol
	50 mM Tris-HCl, pH 7.5
	10 mM $MgCl_2$
	3 mM DTT (added immediately before use)
Buffer A:	10 mM HEPES-KOH, pH 7.9
	1.5 mM $MgCl_2$
	10 mM KCl
	0.5 mM DTT (added immediately before use)

Buffer D:	20 mM HEPES-KOH, pH 7.9
	0.2 mM EDTA
	50 mM KCl
	20% (v/v) Glycerol
	0.5 mM DTT (added immediately before use)

II.3.4.3 SDS polyacrylamide gel electrophoresis (SDS-page)

All SDS polyacrylamide gels were run using two sealed 16 x 16 cm glass plates separated by 1.5 mm spacers and using 14 well combs. The resolving gel solution was prepared, poured between the plates and overlaid with water-saturated butanol. The gel was allowed to set at room temperature, the butanol was washed off with sterile, distilled water, the stacking gel was poured in between the plates and the comb was fixed into place. After polymerisation was complete, the seal was removed from the plates and the comb was gently taken out. The plates were firmly fixed in the "ATTO" electrophoresis apparatus and the chambers filled with 1 x protein gel running buffer. Protein samples were mixed with an equal volume of 2 x SDS loading buffer, heated to $96^{o}C$ for 5 min and centrifuged at 14,000 rpm for 60 seconds. The samples were loaded onto the gel and run at 20-30 mA (or 7 mA overnight) until the bromophenol blue dye front had reached the bottom of the resolving gel.

Broad range pre-stained molecular weight protein markers (6.5-175 kDa) were used (New England Biolabs). Thirty microlitres were loaded.

2 x SDS loading buffer:	100 mM Tris-HCl, pH 6.8
	200 mM DTT (added immediately before use)
	4% (w/v) SDS
	0.2% (w/v) Bromophenol blue
	20% (v/v) Glycerol

10 x protein gel running buffer:

$$250 \text{ mM Tris-base}$$
$$1.9 \text{ M Glycine}$$
$$1\% \text{ (w/v) SDS}$$

12% resolving gel:

16 ml 30% (w/v) Acrylamide/0.8% (w/v) bisacrylamide

13.3 ml Sterile, distilled water

10 ml Resolving gel buffer

60 μl TEMED

200 μl 10% (w/v) Ammonium persulfate

6% stacking gel:

2 ml 30% (w/v) Acrylamide/0.8% (w/v) bisacrylamide

5.5 ml Sterile, distilled water

2.5 ml Stacking gel buffer

6 μl TEMED

150 μl 10% (w/v) Ammonium persulfate

Resolving gel buffer:

1.5 M Tris-HCl, pH 8.8

0.4% (w/v) SDS

Stacking gel buffer:

500 mM Tris-HCl, pH 6.8

0.4% (w/v) SDS

II.3.4.4 Coomassie staining of SDS polyacrylamide protein gels

In order to visualise total protein in an SDS polyacrylamide gel, the gel was incubated in coomassie solution for 30 minutes and then destained by incubation for 7-12 hours in destaining solution under constant gentle shaking at room temperature. The destaining solution was replaced frequently. After destaining, the gel was dried on Whatman 3MM paper using a Hybaid Gel-Vac.

Coomassie solution: 0.1% (w/v) Coomassie blue

50% (v/v) Methanol

10% (v/v) Acetic acid

Destaining solution: 10% (v/v) Methanol

10% (v/v) Acetic acid

II.3.4.5 Western blotting

Electrophoretic transfer of proteins to PVDF immobilion-P membrane - Proteins were transferred electrophoretically from the SDS polyacrylamide gel to Immobilion-P membrane (millipore) using a Bio-rad transfer system and following the manufacturers' protocol. All transfers were performed in 1 x western transfer buffer at 100 V for 90 minutes. Transfer was confirmed by Ponceau S staining and the ponceau stain was removed by washing the membrane for 10 minutes with water and then for 3 minutes in 1 x TBS.

10 x Western transfer buffer: 1.5 M Glycine

200 mM Tris-HCl, pH 8.3

Antibody binding - Non-specific interactions were blocked by incubating the membrane for 60 min (at room temperature) in 80 ml of blocking buffer (3% (w/v) BSA in 1 x TBS (table II.1.3)) with constant shaking. The membrane was then washed twice in 50 ml of 1 x TBS-TT (table II.1.3) and once in 50 ml of 1 x TBS, each time for 10 minutes. Primary antiserum was diluted in 15 ml of fresh blocking buffer (3% BSA (w/v) in 1 x TBS), applied to the membrane and incubated for 2 hours at room temperature. The membrane was then washed again as above. The secondary antibody was applied to the membrane diluted in 10% (w/v) dry milk in 1 x TBS for 1 hour at room temperature. The membrane was then washed four times for 10 minutes in 1 x TBS-TT.

(In experiments in which a protein A epitope was supposed to be detected on the membrane, the membrane was directly incubated with anti-mouse IgG horse radish peroxidase-linked antibody (table II.1.8) diluted in 1 x TBS, 10% (w/v) dry milk after blocking.)

Immunodetection with enhanced chemiluminescence (ECL) - Four millilitres of developer solution was prepared as described by the manufacturer (Amersham), applied to the membrane and incubated for 60 seconds at room temperature. Care was taken to ensure the membrane was uniformly covered by the developer solution. After incubation, the developer solution was removed, the membrane placed in Saran wrap and exposed to autoradiography film. Exposure times were varied depending on the strength of the signal.

Immunodetection with alkaline phosphatase - Five millilitres of NBT/BCIP stain solution were prepared and applied to the membrane. The assay was allowed to continue until the development was considered sufficient and was then stopped by removing the stain solution and washing the membrane in sterile, distilled water.

NBT/BCIP staining solution: 0.5 ml 1 M Tris-HCl, pH 9.5

4.3 ml Sterile, distilled water

125 µl 4 M NaCl

50 µl 0.5 M $MgCl_2$

33 µl NBT (Promega)

16.5 µl BCIP (Promega)

II.3.4.6 *E. coli* expression and purification of a His-tagged yeast protein

Protein expression - Two litres of LB medium (table II.1.1) containing the appropriate antibiotics (section II.1.2.2) were inoculated with 6 ml overnight culture of the expression plasmid-bearing *E. coli* strain. The culture was grown at 30°C to an OD_{600} value of 0.3-0.4. For induction of the expression, IPTG was added to a final

concentration of 0.75 mM and the culture was incubated for 12 hours at 30°C. The cells were harvested by centrifugation at 5000 rpm for 10 minutes at 4°C (Beckman JLA 10.500 rotor).

Preparation of protein extract - The cells were resuspended in 60 ml of Lysis buffer A and 0.2 mg/ml lysozyme was added and the extract was stirred on ice for 45 minutes. After addition of 0.1% (v/v) Triton-X-100 and stirring on ice for another 5 minutes the lysate was transferred to polypropylene centrifuge tubes and centrifuged at 17,000 rpm for 45 minutes at 4°C (Beckman JA 25.50 rotor). The supernate was transferred to a falcon tube and stored at -70°C.

Ni-NTA affinity purification of the His-tagged protein - Two millilitres of Ni-NTA agarose (Qiagen) were washed with 10 ml [20 mM Tris-HCl pH 7.5, 0.5 M NaCl] for 45 minutes at 4°C on a rotating wheel. The Ni-NTA agarose was sedimented by centrifugation at 3,500 rpm for 3 minutes at room temperature (Mistral 1000 centrifuge), the supernate removed and 35 ml of protein extract added. The mixture was incubated for a further 2 hours at 4°C. The Ni-NTA agarose was sedimented by centrifugation at 3,500 rpm for 3 minutes at room temperature (Mistral 1000 centrifuge), the supernate removed and the Ni-NTA agarose resuspended in a residual ml of the mixture and transferred to a small disposable column. All manipulations from now on were done in the cold room at 4°C. The column was washed over 1 hour with 10 ml of BC100 buffer containing 20 mM imidazole (pH 7.9). The flow through was collected as a batch. A second, more stringent wash step involved the addition of 10 ml BC100 buffer containing 55 mM imidazole (pH 7.9). Again, the flow through was collected as a batch. Eventually, the His-tagged protein was eluted with 10 ml of BC100 buffer containing 200 mM imidazole (pH 7.9) over 1 hour and the eluate was collected in fractions of 1.5 ml.

Thirty microlitre aliquots of the eluates and the washes were then resolved by SDS polyacrylamide gel electrophoresis (section II.3.4.3). The total proteins were visualised by coomassie staining (section II.3.4.4) and the His-tagged protein detected by western blotting analysis (section II.3.4.5).

Fractions containing the purified protein were then dialysed against 2 l of BC100 using a Slide-A-Lyzer dialysis cassette (10 kDa cut-off, Pierce).

Lysis buffer A:	50 mM Tris-HCl, pH 7.5
	250 mM NaCl
	10% (w/v) Sucrose

BC100:	20 mM Tris-HCl, pH 7.9
	100 mM KCl
	20% (v/v) Glycerol

Imidazole was prepared fresh immediately before use as a 1 M stock solution (pH adjusted to 7.9 with HCl) and was filtersterilised .

II.3.5 Splicing methods

II.3.5.1 *In vitro* splicing reaction

A splicing reaction mix was prepared in a microcentrifuge tube as follows:

5 x splicing buffer	5 μl
30% (w/v) PEG$_{8000}$	1 μl
Sterile distilled water	1 μl
Splicing extract (section II.3.4.2)	5 μl
Diluted, radiolabelled transcript (section II.3.3.4)	1 μl

The reaction was incubated at 24°C for 25 minutes and terminated by placement on ice. Two microlitres of proteinase K solution was added to the reaction and the sample incubated at 37°C for 45 minutes. After that 100 μl of splicing cocktail was added and the resulting mixture extracted twice with 100 μl of P/C/I (section II.3.1.2) to remove protein debris. The aqueous phase was removed to a fresh microcentrifuge tube and the RNA precipitated by addition of 500 μl of ethanol (abs.), incubation at -70°C for 30 minutes and centrifugation at 14,000 rpm for 20 minutes at 4°C. The RNA pellet was dried and resuspended in 3 μl 3 x loading buffer (section II.3.2.2). The reaction

products were heated to 90°C for 3 minutes before being loaded onto a 6% (w/v) denaturing polyacrylamide gel (section II.3.3.3) at a setting of 24 W for approximately 1 hour and visualised by autoradiography.

5 x splicing buffer:	300 mM KPO_4, pH 7.5
	12.5 mM $MgCl_2$
	10 mM ATP
Proteinase K solution:	1 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
E. coli tRNA:	20 mg/ml *E. coli* tRNA extracted five times with phenol and once with chloroform and precipitated with ethanol.

II.3.5.2 Spliceosomal coimmunoprecipitation analysis

Antibody binding to protein A-sepharose - An appropriate mass of protein A sepharose (8 mg per immunoprecipitation) was hydrated in 1 ml of NTN for 10 minutes and then washed 3 times in 1 ml of NTN. The protein A-sepharose was mixed with 300 μl of NTN and was incubated with an appropriate amount of antibody (table II.1.9) on a rotating wheel overnight at 4°C.

Immunoprecipitation - 60 μl of IgG-agarose and an equivalent amount of agarose beads (without antibody, swollen for 10 minutes in 1 ml of NTN) and the protein A-sepharose with the bound antibody (see above) were treated in parallel as follows:

The samples were washed three times in 1 ml of NTN for each wash and were then incubated in 100 μl blocking solution for 1 hour on a rotating wheel at room temperature. The samples were washed 4 times with 1 ml NTN and once with 1 x PPT. The PPT was taken off and 45 μl of 2 x PPT and 0.5 μl RNasin were added to the samples, which were then mixed with 45 μl splicing extract (section II.3.4.2) (With this extract an *in vitro* splicing reaction , as described in section II.3.3.4, but scaled up 5 times, had been performed, using 1 μl of precipitated radiolabelled (2000 cps/μl) actin transcript as substrate. 5 μl of this splicing reaction was subtracted from the extract after splicing to be deproteinised and run on a polyacrylamide gel (section II.3.5.1) to assay the splicing reaction (total control)). The samples were then incubated on a rotating wheel for 2 hours at 4°C for immunoprecipitation. After that the beads were washed twice with 1 ml of NTN and once with 1 ml of NT. The NT buffer was taken off and 50 μl of proteinase K solution (section II.3.5.1) was added to the beads, which were incubated for 30 minutes at 37°C. The samples were subjected to extraction with P/C/I and ethanol precipitation as described in section II.3.5.1 and run on a 6% polyacrylamide gel. The coprecipitated, labelled RNA species were then visualised by autoradiography.

NTN: 150 mM NaCl
      50 mM Tris-HCl, pH 7.5
      0.1% (v/v) Nonidet P40

NT: 150 mM NaCl
      50 mM Tris-HCl, pH 7.5

2 x PPT: 12 mM HEPES
          300 mM NaCl
          5 mM $MgCl_2$
          0.1% (v/v) Nonidet P40

Blocking buffer:	100 µg/ml BSA
	100 µg/ml Glycogen
	100 µg/ml tRNA

II.3.5.3 Coimmunoprecipitation of snRNAs from protein extracts

Immunoprecipitation - The coimmunoprecipitation of snRNAs was essentially performed as the spliceosomal precipitation (section II.3.5.2) with the exception that the coimmunoprecipitation was done directly with 50 µl neat extract mixed with the antibody-beads in 50 µl 2 x PPT (section II.3.4.2). After immunoprecipitation, washing, deproteinasing, extraction with P/C/I and precipitation of the RNAs, the samples were run on a 6% denaturing polyacrylamide gel as described in section II.3.5.1.

Northern blot analysis of snRNAs - The RNA was transferred electrophoretically to Hybond-N nylon membrane at 60 V for 60 minutes in 0.5 x TBE (table II.1.3) and cross-linked to the membrane as described in section II.3.2.17.

Hybridisation with end-labelled oligonucleotide probes - The membrane was pre-hybridised for 2 hours at 45°C in SES1 buffer (section II.3.2.17). End-labelled oligonucleotide probes (section II.3.2.15) for the U snRNAs (table II.1.6) were mixed and hybridised to the RNA in 20 ml of fresh SES1 buffer overnight at 30°C. (The temperature had to be adjusted to 5°C less than the lowest annealing temperature calculated for the oligonucleotides present.)

Annealing temperature: $4(G + C) + 2(A + T)$.

The membrane was washed five times (1 hour each wash) with SES3 buffer at 30°C and the result visualised by autoradiography.

SES3 buffer: 5% (w/v) SDS

                      1 mM EDTA

                      0.5 M Sodium phosphate buffer, pH 7.2

II.4 Computer analysis

Yeast database searches were performed on the Saccharomyces Genome Database (SGD) network (http://genome-www.stanford.edu/Saccharomyces/).

Protein database searches were performed on the NCBI network server using the BLAST algorithm (Altschul *et al.*, 1990); (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-blast?Jform=0).

Pairwise protein alignments were done using the GENESTREAM ALIGN program on the GENESTREAM network server IGH Montpellier France (http://vega.crbm.cnrs-mop.fr/bin/align-guess.cgi). Multiple protein sequence alignments were performed using the PILEUP program (Pearson and Lipman, 1988) in the GCG9 suite of sequence analysis programs (Devereux *et al.*, 1984). Sequence identities and similarities were identified using the BOXSHADE 3.21 server (http://ulrec3.unil.ch/software/ BOX_form.html).

Chapter III

The Slu7p and Prp18p
two-hybrid screens

III.1 Introduction to the yeast two-hybrid system

The two-hybrid system is a yeast-based genetic assay for detecting protein-protein interactions (Fields and Song, 1989; Chien *et al.*, 1991; Fromont-Racine *et al.*, 1997; Brent and Finlay, 1997 and references therein). The assay takes advantage of the modular organisation of transcriptional activator proteins, which consist of two distinct functional domains, a DNA-binding domain (BD) and a transcription activation domain (AD). The BD localises the transcription factor to the promotor region of a gene, where the AD can then interact with additional components of the transcription machinery to activate transcription. In the yeast two-hybrid assay, the two functional domains are physically separated from each other and become non-covalently brought together again through the interaction of any two proteins to reconstitute a functional transcription factor. The application of the assay requires the co-production of two chimeric fusion proteins with one cell: i) the bait protein consisting of the BD fused to a protein of interest (X); ii) the prey protein consisting of the AD fused to a protein (Y) the interaction of which with protein X is the subject of investigation. If the proteins X and Y interact with each other, a functional transcriptional activator is reconstituted and this can be monitored by successful transcription of one or several reporter genes, which relies upon the functionality of this transcriptional activator (figure III.1). The two-hybrid system can thus be used to investigate the interactions of two proteins or alternatively, by constructing a library of prey fusion proteins, a large number of proteins can be screened to identify proteins which interact with protein X.

The two-hybrid system has advantages over other genetic screening techniques, which in many cases rely on the availability of a mutant phenotype such as screens for suppressor mutants, high copy-number suppressors or screens for synthetic (synergistic) lethal mutations. For non-essential genes or genes for which a conditional phenotype is not available, these techniques are not applicable. In contrast, in the two-hybrid system any wild type protein can be used. In *Saccharomyces cerevisae*, the whole genome has been sequenced (Cherry *et al.*, 1997 and SGD, see chapter II.4) which allows the rapid identification of all interacting factors found in a two-hybrid screen. The rapidly progressing sequencing projects in many organisms

will soon allow the immediate identification of any interacting prey also from libraries of other organisms.

Certainly, the two-hybrid system has its limitations: Especially when using non-yeast proteins, modifications of the proteins required for their function might remain undone and possible interactions will be missed for that reason. Problems might arise when using non-nuclear proteins, which may not fold into their native structure in the nuclear environment. In particular proteins with hydrophobic domains (e.g. integral membrane proteins) will be in many cases unsuitable for use in the two-hybrid system.
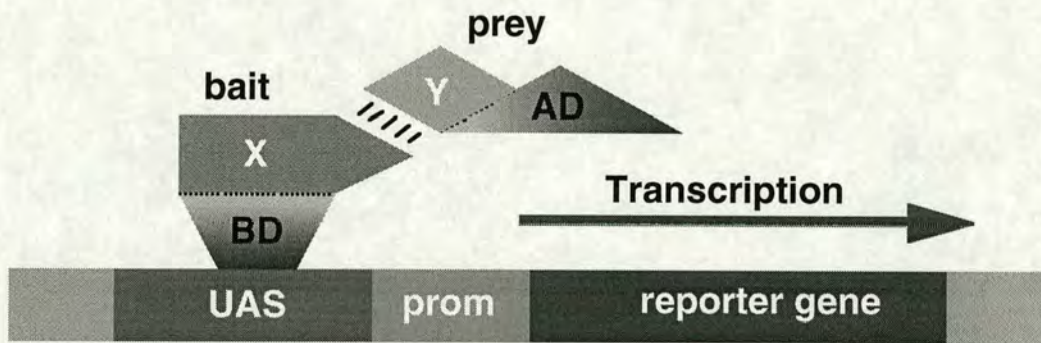
Figure III.1: Basic scheme of the two-hybrid system. Two hybrid proteins are produced within one cell: the bait fusion protein, consisting of the DNA-binding domain (BD) of a transcriptional activator fused to a protein of interest X and the prey fusion, consisting of the transcription activation domain (AD) of a transcription factor fused to a protein Y. The interaction between the proteins X and Y brings the two previously separated functional domains of the transcription factor together and can thus be monitored through the successful transcription of a reporter gene. The BD binds to an upstream activation sequence (UAS) of the reporter gene promotor (prom) and thereby localises the AD, which is then able to promote transcription.

The proteins used as baits in the two-hybrid screens performed in this work were either fused to the BD of the yeast Gal4p transcription factor or to the bacterial LexA protein. The preys were fused to the AD of the Gal4 protein. As reporter genes, the auxotrophic marker gene *HIS3* and/or the *LacZ* gene were used. The fact that the BD and the AD do not need to be derived from the same protein nor from the same

organism illustrates that for the functionality of a transcription factor its localisation is the primary requirement, structural constraints are often secondary.

Some bait proteins do activate the transcription of the reporter genes even in the absence of any prey fusion protein (auto-activating baits). When using the *HIS3* reporter, this problem can be overcome by the addition of the competitive inhibitor 3-aminotriazole (3AT) to the growth medium. This chemical increases the stringency of the selection for histidine prototrophy. Higher levels of *HIS3* transcription (a good transcriptional activator) are required to overcome the histidine auxotrophy of the strain.

In this work, two-hybrid screens were performed following a protocol by Fromont-Racine *et al.* (1997). Also the yeast genomic DNA library (FRYL) used in this work had been generated by the authors. Modifications of the basic two-hybrid assay by the authors allowed a rapid and exhaustive screening of the library. The FRYL comprises 5×10^6 clones, with randomly generated inserts of an average size of 700 basepairs. Given the size of the yeast genome (14×10^6 bp) a fusion event occurs statistically once every 4 base pairs and thus, an in-frame fusion of a genomic DNA fragment with the *GAL4* activation domain sequence can be expected once every 24 nucleotides. Any given ORF should therefore be present in multiple fusion fragments in the FRYL. However due to the nature of the library, whether a given prey will be selected in a screen depends on the size of the interaction domain and on the location of the domain within the protein fragment. Therefore, to evaluate the statistical significance of an interaction all prey proteins are classified into one of 5 categories (figure III.2, Fromont-Racine *et al.*, 1997). The four A categories correspond to fusions beginning with the sequence of an ORF, while the B category relates to fusions of an intergenic region, an anti-sense strand of an ORF, in a non-protein encoding region (rDNA, telomeric DNA, mitochondrial DNA) or in a Ty retrotransposon element. The A1 category presents the most statistically significant interactions. Proteins are classified as A1 prey, if at least two independent overlapping fusion fragments of that protein were identified in a screen. The common region of these overlapping fragments might help to define the interacting domain within the protein. The three other A categories represent proteins, which were identified as only

95

one fusion within a screen, even if this fusion was found multiple times. The A2 category consists of fusions starting close to the initiation codon of an ORF, and within a distance of 150 nucleotides from an upstream in-frame stop codon. These fusions represent n-terminal interaction domains and are expected to be underrepresented in the library due to the in-frame stop-codon which will interrupt translation. The A3 category represents candidates with large coding inserts (greater than 1000 nucleotides). These proteins might need a large interacting domain for a successful two-hybrid response. Again, these preys will be underrepresented in the library, which has an average insert size of only 700 bp. In the A4 category are all preys classified, which do not fall in any of the other categories.

The classification of the preys is only possible due to the reproducible experimental protocol which allows full coverage of the library within one screen. To statistically cover the library 15 million interactions have to be screened. The mating procedure applied typically allowed screening of 20-60 million diploid cells.

In this work both Gal4 as well as LexA baits were used to screen the FRYL. Important differences arise due to the strain background used to carry the baits: For the Gal4 system, the strain CG1945 was used. These cells have a double reporter system, containing both the *HIS3* and the *LacZ* coding sequences under control of the *GAL1* upstream activation sequence (UAS). However, overexpression of some baits in this strain led to a severe flocculation of the cells, which made it unsuitable to use. These problems have not yet been encountered with the L40 strain, which was used as expression strain for LexA baits. Normally, the L40 strain gave a higher mating efficiency than the CG1945 strain, when mated with the library strain Y187. However, the L40 strain is not mutant for the *GAL4* and *GAL80* genes, and therefore, when mated to Y187 cells, the expression of the *LacZ* gene under control of the *GAL1* UAS will become activated. Therefore in two-hybrid screens using the LexA system, only the *HIS3* reporter could be used.

The two-hybrid screens presented in this work were done as part of the TAPIR network (Two-hybrid analyses of proteins involved in RNA metabolism) in which several groups set out to investigate protein-protein interactions in different processes

of RNA metabolism (splicing, polyadenylation, transcription, RNA-transport) by performing numerous exhaustive two-hybrid screens.
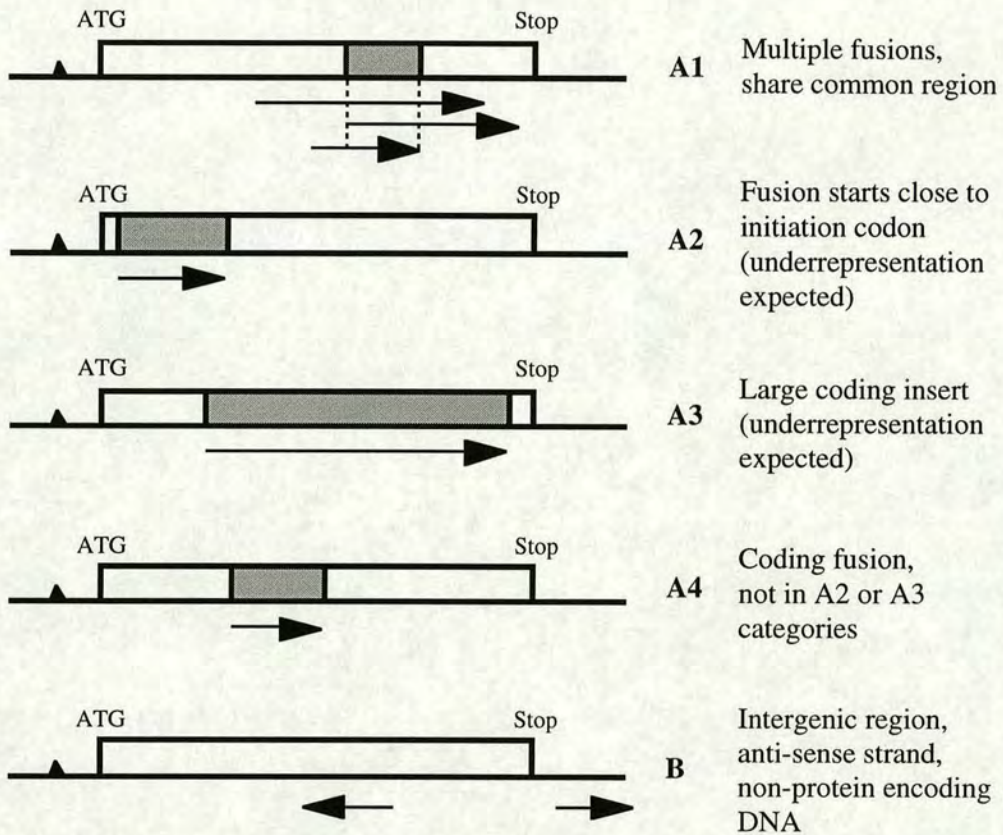
Figure III.2: Classification of prey proteins. Classification of prey identified in a two-hybrid screen of a yeast genomic DNA library. Depicted is a fragment of a yeast chromosome containing an open reading frame (ORF, open box) with start (ATG) and stop codon indicated. The arrows show the orientation and length of prey inserts identified. The shaded boxes represent the region coding for putative interacting domains. A potential in-frame stop codon upstream of the ORF is indicated (black triangle).

III.2 The Slu7p and Prp18p first round screens

First round two-hybrid screens were performed using the full length proteins Slu7p and Prp18p fused to the Gal4p DNA-binding domain (GBD::Slu7p and GBD::Prp18p, respectively) as baits. The *SLU7* and *PRP18* open reading frames (ORFs) had already been cloned into the pAS2ΔΔ vector and transformed into the yeast strain CG1945 (I.Dix and R. van Nues, this laboratory). Two two-hybrid screens using Slu7p as a bait and one two-hybrid screen with the Prp18p as bait had already been done in this laboratory (R. van Nues), so the integrity of the ORFs within the bait plasmids as well as the production of the fusion proteins from the plasmids had been tested. Some of the results of these previous two-hybrid screens will be included in the discussion section below.

III.2.1 The Slu7p two-hybrid screens

Two additional Slu7p two-hybrid screens, using the GBD::Slu7p protein as bait, were performed for mainly two reasons: first, the original library DNA that was used to generate the FRYL had been transformed again into yeast strain Y187 in this laboratory to generate new stock of yeast library. Library derived from this transformation will be referred to as EDYL (Edinburgh yeast library). It seemed to be useful to test this library by performing a two-hybrid screen with a bait which had been used before in order to compare the results and to evaluate the quality of the library. Moreover, it turned out in the previous screens that GBD::Slu7p was a somewhat difficult bait to work with. In the β-gal overlay assay it was challenging to identify putative interactions because generally the blue colour was not very intense, sometimes more grey-greenish than blue. Cells had to be incubated with the toxic X-gal overlay mix for much longer than in screens with other baits in order to develop a blue colour. As a consequence a lot of the putative positive diploids did not survive the overlay assay. Performing additional Slu7p two-hybrid screens, i.e screening much more diploids than actually required to statistically cover the library, should ensure recovery of most positive diploids.

Two-hybrid screen Slu7-III

In two-hybrid screen Slu7-III, cells from 10 vials of EDYL were regenerated and mated with cells from the Slu7p bait culture. A total of 44 million diploid yeast cells were screened for histidine prototrophy. The haploid cells mated with an efficiency of 18.4%. A β-gal overlay assay was performed. After 10 h of incubation 4 blue colonies (clones 1-4) were identified and streaked to colony purify and after 24 h another 17 blue/greenish colonies were picked and streaked (clones 5-21). Three additional very faint blue colonies were identified after 48 h (clones 22-24). After streaking these putative positive clones onto YMM -LWH plates, 13 did not regrow and one was a yeast contaminant present in the library. The remaining 10 positive clones were subjected to a β-gal filterlift assay and the intensity of the blue colour was noted (see table III.1, column "interaction"). The prey plasmids were rescued from these cells and the identity of the inserts was determined by DNA sequencing. The results of the Slu7-III screen are summarized in table III.1 and will be discussed together with the results of the Slu7-IV screen (see below) at the end of this section.

Two-hybrid screen Slu7-IV

In two-hybrid screen Slu7-IV, cells of 1 vial of FRYL were regenerated and mated with cells of the bait culture and 29 million diploid cells were screened for growth on histidine-lacking medium. The mating efficiency was calculated to 24.8%. After 3 days 68 large colonies (clones 1-68) were restreaked onto YMM -LWH plates. These cells had not been subjected to a β-gal-overlay assay in order to optimise survival of putatively positive clones. After 4 days, a large background of smaller colonies had grown. An overlay assay was performed and an additional 72 clones (clones 69-140) that turned blue were restreaked (here 50% more X-gal solution was put into the overlay mix compared to the standard protocol, in order to accelerate development of the blue colour and thus preventing a long incubation time of the cells in the toxic overlay mix). Clones 69-85 were identified as blue after 8 h of overlay assay, clones 86-97 after 13 h and clones 98 -140 after 24h of incubation. Of these 140 clones, 22 did not grow again after restreaking, 4 were contaminants, and only 38 turned blue in the subsequent β-gal filterlift assay. The prey plasmids were rescued from these 38 clones and the identity of the inserts was determined (table III.2).

Table III.1:

Results of two-hybrid screen Slu7-III. A Gal4 DNA-binding domain-Slu7 fusion protein (GBD::Slu7p) was used as bait to screen the EDYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - insert length (nucleotides) determined by sequencing , * insert size has been determined approximately by restriction digest of the prey plasmid with *Bam*HI and subsequent agarose electrophoresis; Interaction - relative strength of response in β-gal filterlift assay: ++ strong response, + moderate response, +/- weak response. Preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	Gene	ORF	Chr	Strand	nt. from AUG	ORF size (bp)	Insert size (bp)	Interaction	Category	Protein info
7-III-19	*NUP170*	YBL079w	II	w	3238	4509	1200*	+/-	A3	Nucleoporin, involved in mRNA export
7-III-5	**PRP22**	**YER013w**	**V**	**w**	**1432**	**3438**	**989**	**+**	**A1**	**Splicing factor**
7-III-9	**PRP22**	**YER013w**	**V**	**w**	**422**	**3438**	**1821**	**+**	**A1**	**Splicing factor**
7-III-16	*DIN1, RNR3*	YIL066w	IX	w	960	2610	800*	+/-	A4	Ribonucleotide reductase, large subunit
7-III-8	*XDJ1*	YLR090w	XII	w	376	1371	800*	++	A4	*E. coli* DnaJ homologue
7-III-24	*SEN1*	YLR430w	XII	w	2950	6696	1400*	+/-	A3	Positive effector of tRNA-splicing endonuclease
7-III-13	*FAP1*	YNL023c	XIV	c	814	2898	2000*	+	A3	Similar to human DNA-binding protein NFX1 (TF)
7-III-4		YNL227c	XIV	c	734	1773	400*	+	A4	Similar to *E. coli* DnaJ protein
7-III-2		YOR275c	XV	c	294	1983	2800*	+	A3	
7-III-1							1100*	++	B	Y'-element, subtelomeric repeat

Table III.2:
Results of two-hybrid screen Slu7-IV. A Gal4 DNA-binding domain-Slu7 fusion protein (GBD::Slu7p) was used as bait to screen the FRYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. No. - frequency with which a fragment was identified in the screen; Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - insert length (nucleotides) determined by sequencing , * insert size has been determined only approximately by restriction digest of the prey plasmid with *Bam*HI and subsequent agarose electrophoresis; Interaction - relative strength of response in β-gal filterlift assay: +++ very strong response; ++ strong response, + moderate response, +/- weak response. Preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size(bp)	Insert size(bp)	Interaction	Category	Protein info
7-IV-44	1	PKC1	YBL105c	II	c	412	3456	500*	+/-	A4	Protein kinase c-like protein
7-IV-3	2	RAD16A	YBR114w	II	w	1664	2373	400*	+	A4	Exision repair protein
7-IV-43	1	LRE1	YCL051w	III	w	480	1761	n.d	+/-	A	Laminarinase resistance protein
7-IV-4	1	MSH6	YDR097c	IV	c	278	3729	1300*	+	A3	Mismatch repair protein
7-IV-131	1	MTH1A	YDR277c	IV	c	401	1302	1000*	+/-	A3	Repressor of hexose transport genes
7-IV-47	1	GLN3	YER040w	V	w	633	2193	1500*	+/-	A3	Transcription factor for pos. nitrogen metabolism
7-IV-48	3	PRP18	YGR006w	VII	w	299	660	900*	++	A4	Second step splicing factor
7-IV-55	1	CBP4	YGR174c	VII	c	-331	513	600*	+/-	A2	Ubiquinol-cyt. c reductase assembly factor
7-IV-32	1	PFK1	YGR240c	VII	c	2328	2964	1950*	+/-	A4	Phosphofructokinase alpha subunit
7-IV-2	2		YGR266c	VII	c	49	2106	500*	+++	A4	Coiled-coil domain containing
7-IV-140	1	MYO1	YHR023w	VIII	w	3724	5553	1400*	+	A3	Myosin heavy chain type II
7-IV-23	**1**	**SSY5**	**YJL156c**	**X**	**c**	**26**	**2064**	**1184**	**++**	**A1**	**Amino acid transport protein**
7-IV-27	**1**	**SSY5**	**YJL156c**	**X**	**c**	**384**	**2064**	**1301**	**++**	**A1**	**Amino acid transport protein**
7-IV-90	1		YLL010c	XII	c	267	1284	1800*	+/-	A3	
7-IV-46	1		YLR320w	XII	w	1849	4365	2100*	+	A3	
7-IV-45n	1	PSE1	YMR308c	XIII	c	1432	3270	n.d	+	A	Protein secretion
7-IV-45c	1		(YLR225c)	XII	c		1224	n.d	+	A	
7-IV-58	1		YOL089c	XV	c	241	3093	800*	+	A4	
7-IV-101	1		YOL091w	XV	w	484	1830	1200*	+/-	A3	Similar to transcription factors, Zn-cluster
7-IV-15	1		YOR191w	XV	w	1336	4860	950*	+/-	A4	
7-IV-75	1		YOR275c	XV	c	324	1986	2472*	+/-	A3	
7-IV-86	2	ADR6, SWI1	YPL016w	XVI	w	1567	3945	1000*	+	A3	Transcription factor
7-IV-50	1	MET31	YPL039w	XVI	w	27	951	1000*	+/-	A3	Regulates sulphur aminoacid metabolism
7-IV-123	1	FAS2	YPL231w	XVI	w	1600	5664	2000*	+	A3	Alpha subunit fatty acid synthase
7-IV-64	1	SPE3	YPR069c	XVI	c	98	882	600*	+	A4	Spermidine synthase
7-IV-63	1	END13	YPR173c	XVI	c	-21	1314	600*	+	A2	Vacuolar sorting, ATPase motif
7-IV-139	1	anti	(YBR140c)	II	w			n.d.	+/-	B	
7-IV-60	1	anti	(YFR020w)	VI	c			n.d.	+/-	B	
7-IV-79	1	anti	(YNL065w)	XIV	c			n.d.	+/-	B	
7-IV-72	1	anti	(YOR195w)	XV	c			n.d.	+/-	B	
7-IV-106	1		Ty-element					n.d.	+/-	B	
7-IV-25	1		Y'-repeat					n.d.	++	B	

In the subsequent discussion(s) of the prey candidates that were found in the Slu7-III, Slu7-IV and any other two-hybrid screens that were performed in the course of this work, the focus will mainly lie on preys that were subjected to further investigation after identification in the screen. Some other candidates will be discussed briefly, if they are of particular interest. A basic description of the protein function of the preys, if known, is given in the tables.

Prey-proteins found in the Slu7p screens

The Slu7p bait was found to interact with two known splicing factors, Prp22p and Prp18p in the screens. Two independent prey fragments of splicing factor Prp22p, a 130 kDa RNA-helicase of the DEAH-box protein family (Schwer and Gross, 1998; de la Cruz *et al.*, 1999), were identified in screen Slu7-III (see figure III.3).

clone:	fragment:		No.
7-III-5:	Prp22p(479-826)		1 x
7-III-9:	Prp22p(142-747)		1 x
Prp22p			

1   479   747   1145

Figure III.3: Prp22p prey fragments identified with the Slu7p bait in two-hybrid screen Slu7-III. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen. A putative RNA-binding motif within Prp22p is indicated (S1, see text).

The common region of the two prey fragments that may be important for the interaction with Slu7p stretches over the aminoacid residues 479-747. This region includes the motifs I to IV (but not V and VI) that are conserved among the members of the DEAD/H-box family. A putative RNA binding motif, similar to that initially found in the bacterial ribosomal protein S1, but now known to be present in a large number of RNA-associated proteins (Gribskov, 1992 and Bycroft *et al.*, 1997) does

not seem to be required for the interaction with Slu7p, because this motif is only contained in one of the fragments identified.

In screen Slu7-IV, the 24 kDa splicing factor Prp18p was identified as an A4 candidate a total of 3 times. The prey fragment consists of the c-terminal 118 amino acid residues of the 219 aminoacid long Prp18 protein (see also figure III.6 B).

The only other A1 candidate found in the Slu7 screens was the 76 kDa Ssy5p, encoded by ORF *YJL156c*. The two prey fragments found in screen Slu7-IV are depicted in figure III.4. The common domain of the fragments stretches from amino acid 129-404 of the 687 amino acid long protein.

clone:	fragment:		No.:
7-IV-23:	Ssy5p(10-402)		1 x
7-IV-27:	Ssy5p(129-560)		1 x

Ssy5p

1   129   402   687
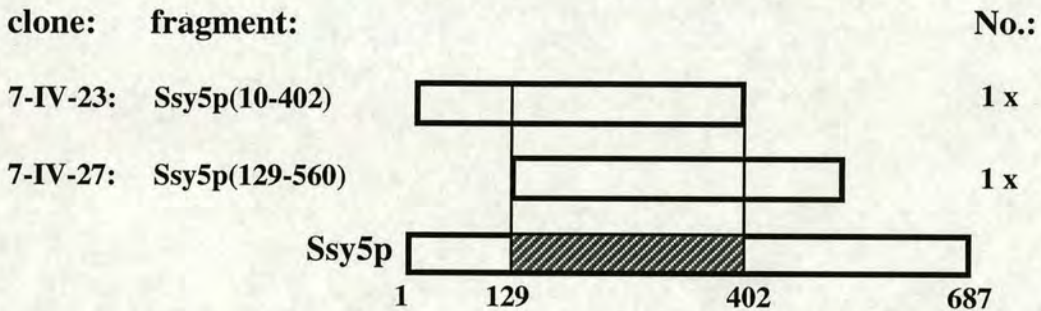
Figure III.4: Ssy5p prey fragments identified with the Slu7p bait in two-hybrid screen Slu7-IV. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen.

The *SSY5* gene was first isolated in a screen for mutants sensitive to the amino acid analog and herbicide **su**lfonylurea and the protein has thus been implicated in amino acid uptake (Joergensen *et al.*, 1998). Nevertheless, the protein was investigated for a putative role in pre-mRNA splicing and the results of this work will be discussed in detail in chapter IV.1.

A protein with a predicted molecular weight of 75 kDa, encoded by ORF *YOR275c*, was found independently in screens Slu7-III and Slu7-IV as A3 clones (see figure III.5).

clone:	fragment:		No.
7-III-2:	Yor275p(99-661)		1 x
7-IV-75:	Yor275p(108-661)		1 x
	Yor275p		

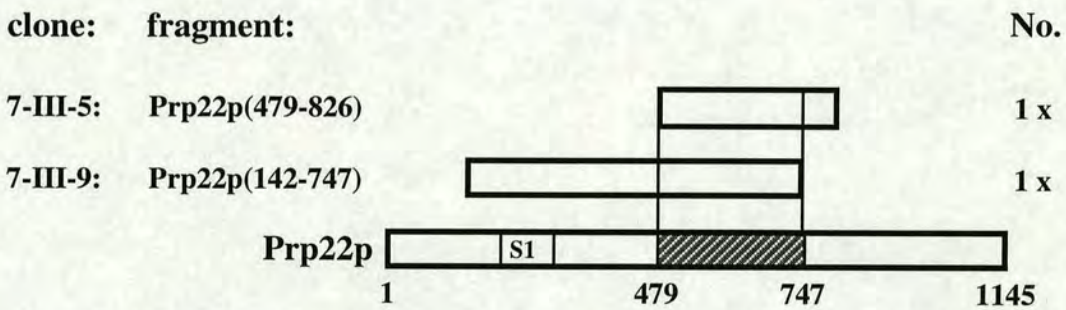1   108   661
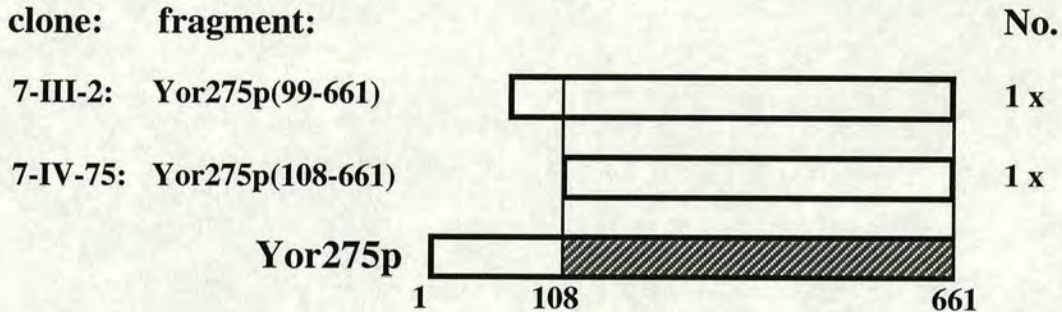
Figure III.5: Yor275p prey fragments identified with the Slu7p bait in two-hybrid screens Slu7-III and Slu7-IV. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen.

The common region of the fragments consists of the amino acid residues 108-661 of the protein. Since the protein was reported to interact also with the splicing factor Snp1p in a two-hybrid screen (Fromont-Racine *et al.*, 1997; Smith and Barrell, 1991), it was chosen for further investigation. A "second round" two-hybrid screen was performed, using Yor275p as a bait. The results of this two-hybrid screen are presented in chapter IV.2.

III.2.2 The Prp18p two-hybrid screen

In the two-hybrid screen using Prp18p as a bait (screen M18), cells of 10 vials of EDYL were mated with the cells of the Prp18p bait culture and 31 million diploid clones were screened for histidine prototrophy. The haploids mated with an efficiency of 20 %. A β-gal overlay assay was performed and 17 blue colonies (clones M18-1 - M18-17) were streaked for colony purification after 6 h of incubation. An additional 31

blue colonies (clones M18-18 - M18-49) were picked and streaked after 24 h incubation with the X-gal overlay mix. Of these 49 putative positive clones, 31 clones survived the overlay assay. The relative strengths of the two-hybrid interactions were then investigated in a filterlift assay. The prey plasmids of these 31 clones were isolated and their identities determined. A summary of the results from this screen is given in Table III.3.

Prey-proteins found with Prp18p

The most striking result of the Prp18p two-hybrid screen (M18) is the finding that Prp18p strongly interacts with three independent Slu7p fragments (see figure III.6 A), which share an overlapping region of 85 amino acids (aa 165-249). This region lies downstream of a Zinc-finger motif (aa 122-135), which is similar to that found in retroviral nucleocapsid proteins (Katz and Jentoft, 1989).

It must be noted that the most frequently isolated fragment, Slu7p(*161-249), is n-terminally preceded by a 194 aminoacid long fragment, aminoacids 341-535, of the Ixr1 protein (see clone M18-5n in table III.3). The 67 kDa Ixr1p is a transcription factor with numerous stretches of glutamin residues and, in addition, it contains two HMG (high mobility group)-box domains, both features of DNA-binding proteins (Brown *et al.*, 1993; McA'Nulty *et al.*, 1996). Indeed, the region preceding the fragment Slu7p(*161-249), almost entirely consists of the two HMG-box domains. Therefore, the question had to be asked, whether the region of amino acids 161-249 of Slu7p alone is sufficient to support an interaction with Prp18p, or whether the strong activation of the *HIS3* and *LacZ* reporter genes was due to a strong affinity of the Ixr1p fragment for the promotor regions of the reporters. Transcription factors are frequently found to cause false positive responses in the two-hybrid assay by autoactivating expression of the reporter genes (e.g. Fromont-Racine *et al.*, 1997). To confirm that the small Slu7p fragment by itself does interact with Prp18p it had to be separated from the Ixr1p fragment and checked for interaction with the Prp18p bait. This test was performed and is described in the next section.

107

Table III.3:

Results of two-hybrid screen M18. A Gal4 DNA-binding domain-Prp18 fusion protein (GBD::Prp18p) was used as bait to screen the EDYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. No. - frequency with which a fragment was identified in the screen; Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - insert length (nucleotides) determined by sequencing , * insert size has been determined approximately by restriction digest of the prey plasmid with *Bam*HI and subsequent agarose electrophoresis; Interaction - relative strength of response in β-gal filterlift assay: +++ very strong response; ++ strong response, + moderate response, +/- weak response. Preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size(bp)	Insert size(bp)	Category	Interaction	Protein info
M18-4	1	*SLU7*	YDR088c	IV	c	-79	1149	920	A1	+++	**Splicing factor**
M18-45	1	*SLU7*	YDR088c	IV	c	492	1149	1390	A1	++	**Splicing factor**
M18-5c	6	*SLU7*	YDR088c	IV	c	480	1149	583+268	A1	+++	**Splicing factor**
M18-3	3		YMR044w	XIII	w	533	1428	654	A4	+	Glutamic acid-rich domains
M18-33	1	BUB2	YMR055c	XIII	c	787	921	900*	A4	+/-	Cell cycle control protein
M18-24	1	VTH2	YHR202w	VIII	w	1513	1809	450*	A4	+	
M18-23c	1		YJL082w	X	w		2196	3400*	A		
M18-46	1		YJL107c	X	c	-2	1164	1200*	A2/A3	+/-	
M18-15	1		YJL222w	X	w	3264	4650	500*	A4	+/-	Putative membrane glycoprotein
M18-5n	6	IXR1	YKL032c	XI	c	1022	1776	583+268	A4	+++	DNA binding protein
M18-27	1	SEN2	YLR105c	XII	c	921	1134	1000*	A4	+/-	Subunit of tRNA-splicing endonuclease
M18-13	1	CMP1	YLR433c	XII	c	1279	1662	435*	A4	+	Calmodulin-binding subunit of protein phosphatase 2B
M18-12	1	MSH2	YOL090w	XV	w	2297	2901	800*	A4	+/-	
M18-6	1	SLK19	YOR195w	XV	w	2114	2466	444	A4	++	Synergistic lethal with kar3, required for mitosis
M18-14	1	ADR6	YPL016w	XVI	w	1637	3945	900*	A4	+/-	Putative transcription factor
M18-49	1	ORC4	YPR162c	XVI	c	1100	1590	1263	A4	+/-	Origin of replication subunit, gene silencing
M18-1	1			XIII	w			n.d	B	n.d	ORC1-RPS1B intergenic
M18-23n	1			XVI	c			3400*	B	++	YPL064c-YPL067c intergenic
M18-36	1	e.g. Ty1-1		e.g I	c			n.d	B	n.d	

A)

clone:	fragment:		No.
M18-4	Slu7p(1-279)		1 x
M18-5c	Slu7p(*161-249)		6 x
M18-45	Slu7p(165-382)		1 x
	Slu7p		

B)

clone:	fragment:		No.
7-IV-48	Prp18p(101-219)		3 x
	Prp18p		
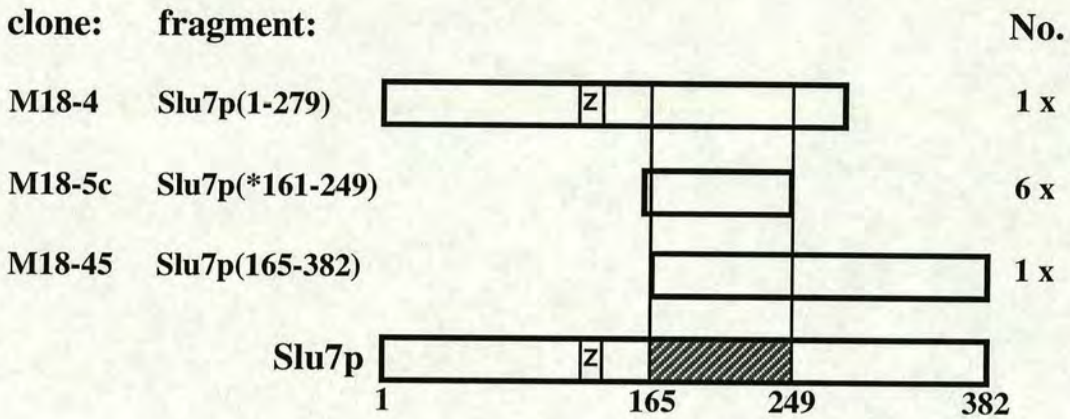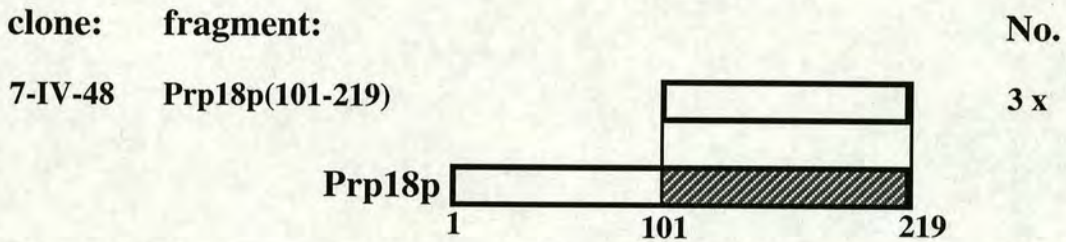
Figure III.6:
A) Slu7p prey fragments identified with the Prp18p bait in two-hybrid screen M18. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen. Z - indicates the position of a Zinc-finger motif (amino acids 122-135). * - the fragment is preceded at the N-terminus by a 194 amino acid long fragment of Ixr1p (see table III.3 and text). **B) Prp18p prey fragment identified with the Slu7p bait in two-hybrid screen Slu7-IV.** Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which the fragment was identified in the screen.

The only additional prey candidate of the Prp18p screen, that was studied in more detail was a protein of unknown function encoded by ORF *YMR044w*. One prey fragment consisting of the amino acid residues 177-395 of the 475 amino acid long Ymr44p was isolated three times. For two reasons the protein was chosen for further

analyses, despite the fact that it was only identified as an A4 candidate: first, Ymr44p was also found in a Prp18 screen performed previously in this laboratory (R. van Nues, personal communication). In that two-hybrid screen it was found also as an A4 interactor, but the fusion was distinct from the one found in screen M18. Second, the protein was found in multiple two-hybrid screens within the TAPIR network with a number of other splicing proteins. These interactions and the investigation of Ymr44p are described in chapter IV.3.

III.2.2.1 The Prp18p-Slu7p interaction

Since the Prp18p-Slu7p two-hybrid interaction was very likely to be of functional relevance, it was important to confirm whether indeed only the small part of Slu7p that was shared by all three prey fragments was able to interact with Prp18p. If, for instance, the fragment Slu7p(*161-249) (see figure III.6) was dependent on the Ixr1p fragment for an interaction, then it would be uncertain, whether the n- or c-terminal regions of fragments Slu7p(1-279) or Slu7p(165-382), respectively, do in fact contain the interacting amino acid residues and not the common domain shared by these fragments.

For this reason, a region spanning the amino acid residues 164-257 of Slu7p (Slu7p(164-257)) was PCR-amplified from plasmid pM18-4 (see table II.1.8) using oligonucleotides 7A and 7B, which contained NcoI and EcoRI restriction sites at their 3' ends, respectively. The region was chosen slightly larger than the actual common region of the prey fragments, because nucleotides at suitable sites, which would guarantee good annealing of the primers, were rare. The PCR product was gel-purified, isolated and subsequently cloned into the pACTIIStop vector (table II.1.7) via the NcoI and EcoRI restriction sites present in the polylinker. The construct was checked by DNA-sequencing using oligonucleotide JC90 (pACTIIStop forward). No PCR-generated deviations from the wild-type sequence were found. The construct, named pMA7c, was then transformed into yeast strain Y187, where it should produce Slu7p(164-257) fused to the Gal4p transcription activation domain (GAD). Subsequently, a directed two-hybrid mating assay was performed to test whether the

94 amino acid long fragment Slu7p(164-257), produced from pMA7c, was able to interact with the Prp18p bait. In order to monitor the strength of the interactions, the cells were spotted onto YMM -LWH plates, containing none, 5, 15, 30 or 50 mM 3AT. The results of this directed two-hybrid assay are shown in figure III.7.

The experiment shows nicely the strong and reciprocal interaction between Slu7p and Prp18p. The interaction of the Slu7p bait with the c-terminus of Prp18p overcomes 3AT concentrations of up to 15 mM. The interaction between Prp18p bait (GBD::Prp18p) and the two-hybrid fusions GAD::Slu7p(*161-249) and GAD::Slu7p(1-279), which were identified in the screen, supports growth of the cells on 3AT concentrations of at least 50 mM. It is important to note, however, that the GAD::Slu7p(164-257) fusion protein also interacts strongly with Prp18p bait (cell growth on 3AT concentrations up to at least 30 mM). This suggests, that the interaction of GAD::Slu7p(*161-249) with Prp18p is not dependent on the Ixr1p fragment, preceding the Slu7p fragment. It seems, though, that the presence of the Ixr1p fragment has some positive influence on the strength of the two-hybrid interaction or maybe on the level of production of the stability of the fusion protein. Taking into account that the common region of the prey fragments only includes residues 165-249, and having shown that the Ixr1 protein fragment is not actually required for the interaction of Slu7p(*161-249) and Prp18p it is tempting to assume that the actual region of interaction within Slu7p does not extend beyond amino acids 165-249.
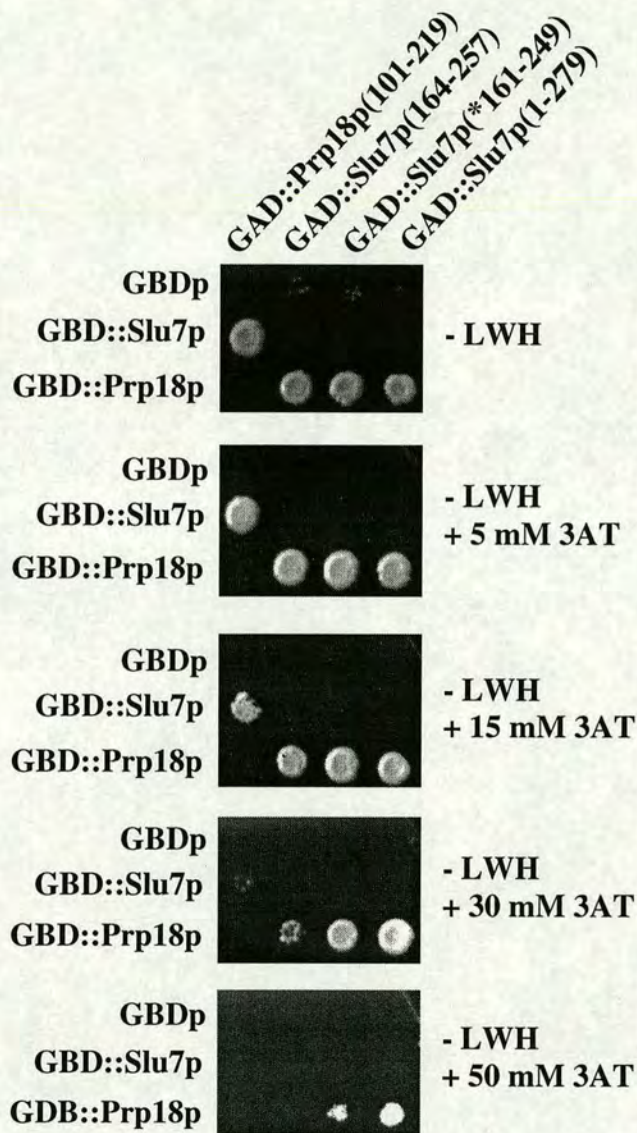
Figure III.7: Investigation of the Slu7p-Prp18p interaction in a directed two-hybrid mating assay. CG1945 and Y187 yeast strains were transformed with pASΔΔ- or pACTIIStop-derived plasmids, respectively, encoding the depicted bait or prey fusion proteins. Transformants were grown on the appropriate selective medium and then mated on YPDA medium overnight at 30°C. The resulting diploids were grown on YMM -WL medium for 24 h at 30°C and transferred onto YMM -LWH selective medium containing different concentrations of 3AT. Cells were incubated for 2 days at 30°C. Bait plasmids used were: GBDp (Gal4p DNA-binding domain), pASΔΔ; GBD::Slu7p, pASSlu7; GBD::Prp18p, pASPrp18. Prey plasmids used were: GAD::Prp18p(101-219), p7-IV-48; GAD::Slu7p(164-257), pMA7c; GAD::Slu7p (*161-249), pM18-5c; GAD::Slu7p(1-279), pM18-4. (GAD - Gal4p transcription activation domain).

113

III.3 Discussion

The performance of two-hybrid screens using the second step splicing factors Slu7p and Prp18p as baits led mainly to two important findings: First, Slu7p and Prp18p interact in a strong and reciprocal way in the two-hybrid screens. Second, Prp22p prey fragments were found to interact with the Slu7p bait.

The Slu7p-Prp18p interaction

The reciprocal two-hybrid interaction between Slu7p and Prp18p is a good confirmation of the previously reported genetic interactions of these second step splicing factors. Mutant alleles of the two proteins had been shown to be synergistic lethal with each other and furthermore, the same alleles were both synthetic lethal with a mutation in the U5 snRNA, suggesting a close functional relationship of the two splicing factors (Frank *et al.*, 1992). In addition, biochemical experiments indicated that the non-essential Prp18 protein somehow facilitates the function of the essential Slu7 protein: in cell extracts that are depleted of Slu7p and Prp18p, splicing is blocked after the first transesterification reaction. Adding back excess Slu7 recombinant protein fully complements the second step block, showing that the requirement of Prp18p for the second step can be obviated by an increased Slu7p activity (Zhang and Schwer, 1997). Adding back Prp18p alone does not restore splicing activity of the doubly depleted extract. Umen and Guthrie (1995a) showed that crosslinking of Slu7p (and Prp8p) to the 3' splice site in a *prp18* mutant extract required prior ATP-hydrolysis by Prp16p and furthermore, that crosslinking is strongest upon addition of Prp18p, suggesting that the function of Prp18p might lie in supporting the association of Slu7p with the 3'-splice site prior to the second transesterification.

The above data suggest a close functional relationship of the Slu7 and Prp18 proteins. However, experiments which proved a direct physical interaction had not been reported. The two-hybrid screen results presented in this work make it tempting to suppose such a direct interaction between the two proteins. First, the two-hybrid interaction is particularly strong, in terms of the intensity of the blue colour in the X-gal overlay- and filterlift assays and in terms of the ability of the diploid cells, coexpressing the interacting fusion proteins, to grow on high concentrations of 3AT.

Second, the region of interaction could be narrowed down to 94 amino acids within Slu7p and to the c-terminal 118 amino acids of Prp18p, which decreases the possibility that other proteins might bridge the interaction. In fact, some potential candidates that could be imagined to bridge an interaction of Slu7p and Prp18p (such as the second step splicing factors Prp22p and Prp16p) have been tested in directed two-hybrid assays for the interaction with the 94 amino acid core fragment Slu7p(164-257) and the c-terminus of Prp18p, but no interacting baits have been found for either of the fragments (data not shown).

While the Slu7p and Prp18p analyses described in this work were ongoing, Zhang and Schwer (1997), obtained results consistent with the data presented above, by performing directed two-hybrid studies with deletion fragments of Slu7p and Prp18p. They were able to further narrow down the region within Slu7p that is required for the interaction and demonstrated that the region from amino acid 200 to 224 is sufficient to support the Slu7p-Prp18p two-hybrid interaction.

The Slu7p-Prp22p interaction

In addition to the interaction with Prp18p, the Slu7p bait interacted with two distinct fragments of Prp22p in screen Slu7-III. Although originally identified as a protein factor required for spliceosome disassembly (Company *et al.*, 1991), recently it was demonstrated that Prp22p in fact has at least one additional function during the splicing process. Schwer and Gross (1998) were able to show in depletion/reconstitution experiments *in vitro* that Prp22p is required in an ATP-independent fashion for the second catalytic step of actin pre-mRNA splicing. Furthermore, they demonstrated, that like Slu7p and Prp18p, Prp22p functions after ATP-hydrolysis by Prp16p. Prp22p protects the 3' splice site against oligonucleotide-directed RNAseH digestion after ATP-hydrolysis by Prp16p, suggesting that it binds to or is at least in close proximity to the 3'-splice site, just like Slu7p and Prp18p. Another feature common to Slu7p, Prp18p and Prp22p is that they are not required for splicing of all introns *in vitro*. Slu7p and Prp18p are dispensable for splicing of precursor RNAs with distances between the branchpoint and the 3'-splice site of 12 or more nucleotides, whereas Prp22p is only needed if the distance exceeds 21 nucleotides (Brys and Schwer, 1996; Zhang and Schwer, 1997; Schwer and Gross,

1998). Therefore it was supposed, that the three proteins are not essential for the general chemistry of the second catalytic step, but that they might act as a molecular bridge between the branchpoint and the 3' splice site, helping to bring the 3' splice site into proximity with the catalytic centre, if the 3' splice site is far away from the branchpoint.

The observed Slu7p-Prp22p two-hybrid interaction gives further evidence for a close proximity and a functionally related action of the two splicing factors and even suggests that Slu7p and Prp22p physically interact with each other in the course of 3' splice site selection and cleavage.

Prp22p was demonstrated to possess an RNA-dependent ATPase and ATP-dependent RNA-helicase activity *in vitro* (Schwer and Gross, 1998; Wagner *et al.*, 1998). The putative interacting domain (the common domain of the two prey-fragments found) with the Slu7p bait spans a region including amino acids 479-747. This region contains motifs I-IV, conserved among the DEAD/H-box family of (putative) RNA-helicases. Motifs I and II are thought to be responsible for ATP-binding and hydrolysis, while motif III has been implicated in the helicase or unwinding function of the DEAD/H-box proteins (Pause and Sonenberg, 1992). It is therefore intriguing to suppose that Slu7p could directly modulate Prp22p action by interacting with these important functional domains. However, it has to be pointed out, that there is no proof that the interaction between Slu7p and Prp22p happens prior to or during the second step. It could well be, that what we see is an interaction of the two proteins at a stage of spliceosome disassembly. It has been shown previously, that Slu7p remains bound to the excised intron and to the mature message until the spliceosome disassembles, in a reaction that requires ATP (Brys and Schwer, 1996). Therefore, the two proteins might be in proximity, when Prp22p acts to disassemble the spliceosome. The finding that Slu7p interacts with domains which are important for the RNA-unwinding function of Prp22p, could support the latter model.

Finally it has to be pointed out that the observed two-hybrid interaction might not be direct. Indeed, compared to e.g. the interaction of Slu7p with Prp18p, the interaction of the Slu7p bait with the Prp22p prey-fragments is relatively weak. In a direct mating two-hybrid assay the diploids grow on histidine-lacking medium, but the

addition of only 5 mM 3-AT to the plates does not allow sufficient expression of the *HIS3*-reporter gene in the cells to overcome the more stringent selection conditions (data not shown). Although a weak two-hybrid interaction does not necessarily mean that the interaction is indirect (as a strong interaction does not prove a direct one), it is reasonable to ask whether there are potential candidates that might be bridging the observed two-hybrid interaction. One such candidate should at least be mentioned briefly here. In a two-hybrid screen with the Prp16p bait performed in this laboratory, Prp22p was found as the only interacting prey protein (Alan Colley, personal communication). Two independent prey-fragments were isolated, that overlap with the Prp22p prey-fragments found with the Slu7p bait (see figure III.8). The common region of these fragments contains the motifs IV and V conserved among the members of the DEAD/H-box protein family, thus sharing a region including motif IV with the common region of the prey-fragments found with Slu7p.
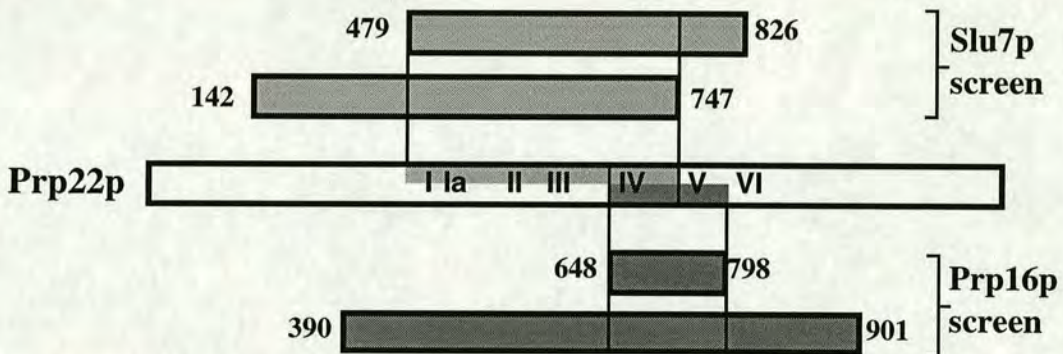
Figure III.8: Putative interacting regions of Prp22p with Slu7p and Prp16p. Depicted are Prp22p prey fragments identified in two-hybrid screens using Slu7p as bait (light grey boxes) or using Prp16p as bait (dark grey boxes). Both baits interact with regions within the helicase domain of Prp22p, which contains conserved sequence motifs (the position of these motifs I-VI are indicated; see text for information on the motifs). The Prp16p two-hybrid screen was performed by A. Colley, this laboratory.

The reporter response produced by the Prp16p-Prp22p two-hybrid interaction is much stronger than the one observed for the Slu7p-Prp22p interaction, withstanding 3-AT concentrations of at least 10 mM (A. Colley, personal communication). This makes a triple ("sandwich") interaction imaginable, with Prp16p being in between Slu7p and

Prp22p. However, a two-hybrid interaction of Prp16p bait with Slu7p prey-fragments was not observed, neither in the Prp16p two-hybrid screen, nor in directed two-hybrid mating experiments. From the two-hybrid data, another situation is more likely, in which Prp22p sits in between Slu7p and Prp16p, or alternatively, in which Slu7p and Prp16p bind side by side to Prp22p. But again, one has to bear in mind, that the two-hybrid data do not give temporal information on the observed interactions. It is likely, that the Prp16p-Prp22p interaction happens immediately prior to the second step, because Prp16p is not detected in the spliceosome at other times. As mentioned earlier, for Slu7p and Prp22p it can not be predicted, at what stage of the splicing reaction the two proteins might contact each other.

Other prey-candidates found in the Slu7p- and Prp18p two-hybrid screens

Two other potentially interesting candidates were found in two independent fusions in the Slu7p screens: For both the Yor275p and Ssy5p not much information was available. Therefore, these candidates were studied in more detail as described in chapters IV.1 and IV.2.

In two-hybrid screen Prp18M, apart from Slu7p no other candidates were found in more than one fusion. Only Ymr44p, was found to be interesting for further studies, since it had been found also in a previously performed Prp18p screen (R. van Nues, personal communication). The investigation of this protein is described in chapter IV.3.

Comparison of two-hybrid screens Slu7-III and Slu7-IV

As mentioned in the introduction of this chapter, the two Slu7p two-hybrid screens were performed with different yeast libraries, the FRYL and the EDYL, which were generated from the same DNA stock but in separate transformations. Although in both screens a sufficient amount of diploid yeast cells was screened to cover the library, very little overlap between the screens was observed. Only one candidate, Yor275p, was found in both screens. This need not necessarily be due to a different quality of the libraries, but could be a consequence of the behaviour of the Slu7p bait. Looking at the tables III.1 and III.2, it can be seen that in both screens, a large collection of candidates was found only once in the screens and most of them produced

only a weak response in the filterlift assay. Potentially interesting candidates, which are likely to be biologically significant interactors, were also not found very often in the screen. Prp18p was only found as one prey-fusion three times, Prp22p only as two independent fusions and these fragments only once. The other A1 prey, Ssy5p, was also only found twice. In most other two-hybrid screens, A1 candidates are represented in more than two fusions and furthermore, the ratio between the total number of identified fusions within the A1 category and fusions of the other categories is generally much higher (own observation). This suggests, that the Slu7p bait is not ideal for use in this system, since it does not seem to function as a good component of a reconstituted transcriptional activator, even in cases in which a strong interaction with the prey could be expected. This behaviour means that it will be difficult with this bait to distinguish between "good or real" two-hybrid interactions among a large background of obviously random, aspecific interactions. With the Slu7p bait it is therefore difficult to predict qualitative differences between the EDYL and the FRYL. Further screens with the same baits need to be done using both libraries to evaluate whether there are major differences with respect to the number of prey fusions present in these libraries.

Chapter IV

Characterisation of prey proteins found in the Slu7p and Prp18p two-hybrid screens

IV.1 Characterisation of the Ssy5 protein

The protein Ssy5 (encoded by the *YJL156c* ORF) was found as the only A1 prey in the two-hybrid screen Slu7-IV. Two overlapping fragments were isolated that interact with the Slu7 protein (see figure III.4, section III.2.2). The protein has a predicted molecular weight of 76 kDa and does not share any significant homology with any proteins in the databases. The *YJL156c* gene was originally identified in a screen for mutants that showed <u>s</u>ensitivity to a <u>s</u>ulfon<u>y</u>lurea herbicide (MM) on rich medium (Joergensen *et al.*, 1998). Cells growing on medium containing MM are dependent on uptake of the amino acids isoleucine and valine from the medium, since this herbicide is a potent inhibitor of the biosynthesis of branched-chain amino acids. Although the above finding suggested an involvement of the Ssy5 protein in amino acid uptake, a more detailed investigation of a possible function of the protein in pre-mRNA splicing was undertaken in this work, in order to clarify the meaning of the observed interaction of the protein with the Slu7p bait.

The Ssy5 protein was also found as an A1 prey candidate in a two-hybrid screen with the splicing factor Isy1p, which had been recently characterised in this laboratory as a non-essential, spliceosome-associated protein (I. Dix, personal communication and Dix *et al.*, 1999). This observation made a further investigation of Ssy5p even more intriguing.

IV.1.1 Deletion of the *SSY5* open reading frame

In order to study the function of the Ssy5 protein a deletion mutant was generated. Using the oligonucleotides 156-1 and 156-2 the *HIS3* marker gene was amplified by PCR from the vector YIp1 and the purified PCR product was transformed into the diploid yeast strain BMA38, which is deleted for the *HIS3* locus. Due to the choice of the primers, the generated PCR product contains 5' and 3'-ends homologous to regions just upstream and downstream of the *YJL156c* ORF respectively, so that integration of the *HIS3* gene at the *YJL156c* locus can occur by homologous recombination, thereby replacing one allele of the *SSY5* gene on the chromosome.

Histidine prototrophic transformants were then tested for integration of the *HIS3* gene at the desired locus by southern blot analysis, as shown in figure IV.1. The diploid strain deleted for one of the *SSY5* alleles is referred to as YMA156/1.

Figure IV.1:
Replacement of the *SSY5* open reading frame with the *HIS3* marker gene on the chromosome. A schematic representation of the *SSY5* wild-type locus as well as of the locus after insertion of the *HIS3* marker gene is given and *Hind*III restriction sites are indicated (H). The region complemantary to the probe used in the southern blot is also shown (P). Genomic DNA was isolated from two independent transformants of strain YMA156/1 as well as from the parental wild-type strain BMA38. The DNA was digested with *Hind*III and resolved in a 1% (w/v) agarose gel and blotted to Hybond-N nylon membrane (Amersham). The blot was probed with a radiolabelled DNA fragment and the result visualised by autoradiography. The black arrow indicates the position of the fragment from the wild-type locus, the white arrow the position of the fragment from the locus after replacement of the *SSY5* gene by the *HIS3* gene.

Subsequently, strain YMA156/1 was sporulated and the formed tetrads were dissected onto YPDA plates (figure IV.2 A). In seven cases all four spores germinated, leading to two small and two larger colonies for each dissected tetrad. In the other three cases, only three spores germinated and again one or two colonies displayed a slow growth phenotype. By streaking out the resultant haploid colonies onto YMM -H solid

122

medium it could be shown, that the *HIS3* marker gene cosegregates with the slow growth phenotype (data not shown), showing that the growth defect results from the deletion of the *SSY5* ORF. When cell growth of the haploid strain (designated as YMA156/2) deleted for the *ssy5* locus was tested at a range of temperatures on solid rich medium, it turned out that the growth defect upon deletion of the *SSY5* gene is exacerbated at higher temperatures, leading to a severe temperature sensitivity at 37°C (figure IV.2 B). These data demonstrate that the Ssy5 protein is dispensable for cell viability, but that the protein is required for normal cell growth on rich medium.

A)

B)

Figure IV.2:
Analyses of effects of *SSY5* deletion from the chromosome.
A) Tetrad analysis of strain YMA156/1. Strain YMA156/1 was grown overnight at 30°C in YPDA, the cells collected and transferred to SPM sporulation medium. After 5 days of incubation at 23°C the formed tetrads were dissected onto YPD agar plates. The spores were incubated for four days at 23°C.
B) Haploid cells deleted for the *SSY5* locus (strain YMA156/2) as well as cells from the parental wild-type strain BMA38 were suspended in dilutions in microtiter plates, spotted onto YPDA agar plates and incubated at a range of temperatures for 2 days.

IV.1.2 Effect of *SSY5* deletion on pre-mRNA splicing *in vivo*

In order to investigate a putative involvement of Ssy5p in pre-mRNA splicing, cells of wt strain BMA38n and of strain YMA156/2 were grown in rich liquid medium at 23°C and were then transferred to prewarmed medium of either 23°C or 37°C. Samples of these cultures were taken after 2 and 5 hours of incubation and total RNA was isolated from these cells. Processing of pre-mRNA was then studied by Northern-analysis, using a radiolabelled probe for exon 1 of the *RP28* RNA (figure IV.3).

Figure IV.3:
Northern analysis of the splicing capability of strain YMA156/2. YMA156/2 cells, deleted of *SSY5*, as well as BMA38n wild-type cells were grown in YPDA medium at 23°C, sedimented by centrifuging at room temperature for 5 minutes and then transferred to pre-warmed YPDA medium at 23°C or 37°C. 10 ml samples of the cultures were taken after different incubation times and total RNA was extracted from the cells. 10 μg of RNA was resuspended in formamide loading buffer, resolved in a 1.2% (w/v) agarose gel and blotted to Hybond-N membrane (Amersham). The blot was then probed with a radiollabelled DNA fragment, complementary to exon 1 of the *RP28* gene. The result was visualised by autoradiography. (YMA45/2 - RNA isolated from strain YMA45/2 grown under repressing conditions was used as control; this strain was known to be splicing deficient under the conditions used). The positions of the *RP28* pre-mRNA and mRNA are indicated.

No obvious splicing defect due to the deletion of the *SSY5* ORF can be detected in this assay. Neither at 23°C, nor at temperatures at which the growth defect is pronounced, was an accumulation of pre-mRNA visible in strain YMA156/2. Strain YMA45/2 (lane 6) serves as a positive control. This strain has a splicing defect and thus accumulates unprocessed precursor RNA (upper band). Furthermore, the temperature shift does not lead to a significant decrease in the amount of mature message produced, supporting the idea that the exacerbated growth defect of strain YMA156/2 at higher temperatures is not a consequence of a pronounced defect in the splicing machinery.

In figure IV.3 generally lower levels of mRNA are detected in the lanes from strain YMA156/2 compared to the wt strain BMA38. However, it has to be pointed out, that these differences are not likely to be due to defects in the splicing machinery. By looking at the ethidium bromide stained total RNA on the northern membrane it becomes clear that less RNA is present in the lanes 6-11 compared to lanes 1-5 (data not shown).

IV.1.3 Effect of *SSY5* deletion on pre-mRNA splicing *in vitro*

It was then tested whether the Ssy5 protein is required for splicing of actin precursor-mRNA *in vitro*. For this purpose, whole cell extracts (splicing extracts) were prepared from cells of strains YMA156/2 (Δ*ssy5*) and BMA38n (wt) grown at 23°C for 12 hours. Splicing was performed at 23°C for 30 min with these extracts after addition of splicing buffer, ATP and labelled actin pre-mRNA. In a parallel sample, extract from strain YMA156/2 was preincubated at 37°C for 25 min, before splicing was performed at 23°C. The result of this experiment is depicted in figure IV.4.

It is clearly demonstrated that the Ssy5p is dispensable for splicing of actin precursor-mRNA *in vitro*, because the cell extract from the *SSY5* deleted strain YMA156/2 splices the substrate pre-mRNA as efficiently as the wt extract. In the wt strain as well as in YMA156/2, both intermediate lariat-intron exon 2 and the excised intron are equally efficient produced. Preheating the YMA156/2-extract to 37°C for 25 min before the splicing reaction is performed does not have a negative effect on the

splicing capability of the extract. This again indicates, as for the *in vivo* result, that the exacerbation of the growth defect of the YMA156/2 strain at higher temperatures is not a consequence of a defect within the splicing machinery that becomes more pronounced at elevated temperatures.

Figure IV.4:
In vitro **splicing analysis of strain YMA156/2.** Whole cell extracts (splicing extracts) were prepared from cultures of strain YMA156/2, deleted for *SSY5*, as well as from BMA38n grown at 23°C for 12 hours to an OD_{600} of 0.5. Splicing reactions were then performed at 23°C for 25 minutes using 5 µl aliquots of the extracts and adding radiolabelled actin pre-mRNA and splicing buffer containing ATP. Alternatively, YMA156/2 extract was heat-treated (h.i. +) by pre-incubation at 37°C for 25 minutes before the splicing reaction was set up. The samples were then deproteinised, the RNA precipitated and loaded onto a 6% (w/v) SDS polyacrylamide gel. The result was visualised by autoradiography. The positions of the pre-mRNA, the lariat-intron and the lariat-intron exon 2 intermediate are indicated.

IV.1.4 Discussion

The experiments described above strongly suggest that the Ssy5 protein, which was found as an A1 prey in a Slu7p two-hybrid screen, is not required for pre-mRNA splicing. Deletion of the *SSY5* ORF from the genome leads to a slow growth phenotype that becomes more pronounced at higher temperatures and leads to a severe temperature sensitivity at 37°C. For many splicing factors it has been shown that mutations in (or deletions of) genes encoding these proteins lead to temperature-sensitive growth phenotypes of the cells. In northern blot analyses of RNAs isolated from such mutant cells an accumulation of pre-mRNA is normally evident already two hours after shift to the non-permissive temperature (e.g. Maddock *et al.*, 1996, Vijayraghavan *et al.*, 1989). For the *SSY5* deleted strain YMA156/2, no pre-mRNA was detected after 5 hours incubation at 37°C, suggesting that the pre-mRNA was efficiently spliced even under non-permissive conditions. It has to be noted, that the YMA156/2 strain does not grow at all in rich liquid medium at 37°C (data not shown). From the *in vivo* data it could not be entirely ruled out, that Ssy5p might be a splicing factor required only for the second transesterification reaction to occur. If this was the case, one might not see an accumulation of pre-mRNA, because it would be efficiently processed into the intermediates, which are not detected in the northern analysis, since a probe for exon 1 was used. Due to the obvious loading differences of the samples a definite statement about the effect on the mRNA levels was also difficult to make. However, the result of the *in vitro* splicing assay strongly argues against this interpretation. Splicing intermediates do not accumulate in the absence of Ssy5p, and the products of the splicing reaction are produced to wt levels.

During the course of the experiments described above, some more information on the function of the Ssy5 protein became available from other work: after the initial identification of the *ssy5* gene in a screen for mutants sensitive to the sulfonylurea-herbicide MM, an inhibitor of the biosynthesis of branched-chain amino acids, the gene was cloned via complementation from a low copy number library (Joergensen *et al.*, 1998). In more directed experiments it was shown that the *ssy5* mutant strain was

deficient in the uptake of branched-chain amino acids such as isoleucine and valine, but that in addition the uptake of several other neutral amino acids, especially leucine and phenylalanine, was effected. If toxic analogs of these amino acids were added to the growth medium, the mutant strain showed resistance to these molecules, which gives further evidence, that amino acid uptake is reduced.

The structural gene for the branched-chain amino acids permease had been identified by Grauslund *et al.* (1995) as *BAP2*. Truncation of Bap2p at its carboxy-terminus leads to an **in**creased uptake of branched-chain amino acids. Joergensen and collaborators found, that the *ssy5* mutation is epistatic over the dominant Bap2p truncation, and they propose that Ssy5p might be involved in regulating the expression of *BAP2*.

If the Ssy5 protein is not involved in the splicing pathway, the question remains why it was found as interactor with two known splicing factors in the two-hybrid screens? The Ssy5 prey fusion proteins identified with the Slu7p- and the Isy1p bait do overlap to a large extend and they share a common region consisting of the amino acid residues 129-238. The region contains 33% charged residues, which might confer some nucleic acid binding capacity. Maybe this gives rise to an aspecific interaction with an RNA molecule that in turn interacts with Slu7p and/or Isy1p. If one thinks of a role of the Ssy5 protein as a regulator of *BAP2* expression, a nucleic acid binding activity could very well be imagined.

IV.1 Characterisation of the Ssy5 protein

The protein Ssy5 (encoded by the *YJL156c* ORF) was found as the only A1 prey in the two-hybrid screen Slu7-IV. Two overlapping fragments were isolated that interact with the Slu7 protein (see figure III.4, section III.2.2). The protein has a predicted molecular weight of 76 kDa and does not share any significant homology with any proteins in the databases. The *YJL156c* gene was originally identified in a screen for mutants that showed <u>s</u>ensitivity to a <u>s</u>ulfon<u>y</u>lurea herbicide (MM) on rich medium (Joergensen *et al.*, 1998). Cells growing on medium containing MM are dependent on uptake of the amino acids isoleucine and valine from the medium, since this herbicide is a potent inhibitor of the biosynthesis of branched-chain amino acids. Although the above finding suggested an involvement of the Ssy5 protein in amino acid uptake, a more detailed investigation of a possible function of the protein in pre-mRNA splicing was undertaken in this work, in order to clarify the meaning of the observed interaction of the protein with the Slu7p bait.

The Ssy5 protein was also found as an A1 prey candidate in a two-hybrid screen with the splicing factor Isy1p, which had been recently characterised in this laboratory as a non-essential, spliceosome-associated protein (I. Dix, personal communication and Dix *et al.*, 1999). This observation made a further investigation of Ssy5p even more intriguing.

IV.1.1 Deletion of the *SSY5* open reading frame

In order to study the function of the Ssy5 protein a deletion mutant was generated. Using the oligonucleotides 156-1 and 156-2 the *HIS3* marker gene was amplified by PCR from the vector YIp1 and the purified PCR product was transformed into the diploid yeast strain BMA38, which is deleted for the *HIS3* locus. Due to the choice of the primers, the generated PCR product contains 5' and 3'-ends homologous to regions just upstream and downstream of the *YJL156c* ORF respectively, so that integration of the *HIS3* gene at the *YJL156c* locus can occur by homologous recombination, thereby replacing one allele of the *SSY5* gene on the chromosome.

Histidine prototrophic transformants were then tested for integration of the *HIS3* gene at the desired locus by southern blot analysis, as shown in figure IV.1. The diploid strain deleted for one of the *SSY5* alleles is referred to as YMA156/1.

Figure IV.1:
Replacement of the *SSY5* open reading frame with the *HIS3* marker gene on the chromosome. A schematic representation of the *SSY5* wild-type locus as well as of the locus after insertion of the *HIS3* marker gene is given and *Hind*III restriction sites are indicated (H). The region complemantary to the probe used in the southern blot is also shown (P). Genomic DNA was isolated from two independent transformants of strain YMA156/1 as well as from the parental wild-type strain BMA38. The DNA was digested with *Hind*III and resolved in a 1% (w/v) agarose gel and blotted to Hybond-N nylon membrane (Amersham). The blot was probed with a radiolabelled DNA fragment and the result visualised by autoradiography. The black arrow indicates the position of the fragment from the wild-type locus, the white arrow the position of the fragment from the locus after replacement of the *SSY5* gene by the *HIS3* gene.

Subsequently, strain YMA156/1 was sporulated and the formed tetrads were dissected onto YPDA plates (figure IV.2 A). In seven cases all four spores germinated, leading to two small and two larger colonies for each dissected tetrad. In the other three cases, only three spores germinated and again one or two colonies displayed a slow growth phenotype. By streaking out the resultant haploid colonies onto YMM -H solid

122

medium it could be shown, that the *HIS3* marker gene cosegregates with the slow growth phenotype (data not shown), showing that the growth defect results from the deletion of the *SSY5* ORF. When cell growth of the haploid strain (designated as YMA156/2) deleted for the *ssy5* locus was tested at a range of temperatures on solid rich medium, it turned out that the growth defect upon deletion of the *SSY5* gene is exacerbated at higher temperatures, leading to a severe temperature sensitivity at 37°C (figure IV.2 B). These data demonstrate that the Ssy5 protein is dispensable for cell viability, but that the protein is required for normal cell growth on rich medium.

A)

B)

Figure IV.2:
Analyses of effects of *SSY5* deletion from the chromosome.
A) Tetrad analysis of strain YMA156/1. Strain YMA156/1 was grown overnight at 30°C in YPDA, the cells collected and transferred to SPM sporulation medium. After 5 days of incubation at 23°C the formed tetrads were dissected onto YPD agar plates. The spores were incubated for four days at 23°C.
B) Haploid cells deleted for the *SSY5* locus (strain YMA156/2) as well as cells from the parental wild-type strain BMA38 were suspended in dilutions in microtiter plates, spotted onto YPDA agar plates and incubated at a range of temperatures for 2 days.

IV.1.2 Effect of *SSY5* deletion on pre-mRNA splicing *in vivo*

In order to investigate a putative involvement of Ssy5p in pre-mRNA splicing, cells of wt strain BMA38n and of strain YMA156/2 were grown in rich liquid medium at $23^\circ C$ and were then transferred to prewarmed medium of either $23^\circ C$ or $37^\circ C$. Samples of these cultures were taken after 2 and 5 hours of incubation and total RNA was isolated from these cells. Processing of pre-mRNA was then studied by Northern-analysis, using a radiolabelled probe for exon 1 of the *RP28* RNA (figure IV.3).

Figure IV.3:
Northern analysis of the splicing capability of strain YMA156/2.
YMA156/2 cells, deleted of *SSY5*, as well as BMA38n wild-type cells were grown in YPDA medium at $23^\circ C$, sedimented by centrifuging at room temperature for 5 minutes and then transferred to pre-warmed YPDA medium at $23^\circ C$ or $37^\circ C$. 10 ml samples of the cultures were taken after different incubation times and total RNA was extracted from the cells. 10 μg of RNA was resuspended in formamide loading buffer, resolved in a 1.2% (w/v) agarose gel and blotted to Hybond-N membrane (Amersham). The blot was then probed with a radiollabelled DNA fragment, complementary to exon 1 of the *RP28* gene. The result was visualised by autoradiography. (YMA45/2 - RNA isolated from strain YMA45/2 grown under repressing conditions was used as control; this strain was known to be splicing deficient under the conditions used). The positions of the *RP28* pre-mRNA and mRNA are indicated.

No obvious splicing defect due to the deletion of the *SSY5* ORF can be detected in this assay. Neither at 23°C, nor at temperatures at which the growth defect is pronounced, was an accumulation of pre-mRNA visible in strain YMA156/2. Strain YMA45/2 (lane 6) serves as a positive control. This strain has a splicing defect and thus accumulates unprocessed precursor RNA (upper band). Furthermore, the temperature shift does not lead to a significant decrease in the amount of mature message produced, supporting the idea that the exacerbated growth defect of strain YMA156/2 at higher temperatures is not a consequence of a pronounced defect in the splicing machinery.

In figure IV.3 generally lower levels of mRNA are detected in the lanes from strain YMA156/2 compared to the wt strain BMA38. However, it has to be pointed out, that these differences are not likely to be due to defects in the splicing machinery. By looking at the ethidium bromide stained total RNA on the northern membrane it becomes clear that less RNA is present in the lanes 6-11 compared to lanes 1-5 (data not shown).

IV.1.3 Effect of *SSY5* deletion on pre-mRNA splicing *in vitro*

It was then tested whether the Ssy5 protein is required for splicing of actin precursor-mRNA *in vitro*. For this purpose, whole cell extracts (splicing extracts) were prepared from cells of strains YMA156/2 (Δ*ssy5*) and BMA38n (wt) grown at 23°C for 12 hours. Splicing was performed at 23°C for 30 min with these extracts after addition of splicing buffer, ATP and labelled actin pre-mRNA. In a parallel sample, extract from strain YMA156/2 was preincubated at 37°C for 25 min, before splicing was performed at 23°C. The result of this experiment is depicted in figure IV.4.

It is clearly demonstrated that the Ssy5p is dispensable for splicing of actin precursor-mRNA *in vitro*, because the cell extract from the *SSY5* deleted strain YMA156/2 splices the substrate pre-mRNA as efficiently as the wt extract. In the wt strain as well as in YMA156/2, both intermediate lariat-intron exon 2 and the excised intron are equally efficient produced. Preheating the YMA156/2-extract to 37°C for 25 min before the splicing reaction is performed does not have a negative effect on the

splicing capability of the extract. This again indicates, as for the *in vivo* result, that the exacerbation of the growth defect of the YMA156/2 strain at higher temperatures is not a consequence of a defect within the splicing machinery that becomes more pronounced at elevated temperatures.

Figure IV.4:
In vitro **splicing analysis of strain YMA156/2.** Whole cell extracts (splicing extracts) were prepared from cultures of strain YMA156/2, deleted for *SSY5*, as well as from BMA38n grown at 23°C for 12 hours to an OD_{600} of 0.5. Splicing reactions were then performed at 23°C for 25 minutes using 5 μl aliquots of the extracts and adding radiolabelled actin pre-mRNA and splicing buffer containing ATP. Alternatively, YMA156/2 extract was heat-treated (h.i. +) by pre-incubation at 37°C for 25 minutes before the splicing reaction was set up. The samples were then deproteinised, the RNA precipitated and loaded onto a 6% (w/v) SDS polyacrylamide gel. The result was visualised by autoradiography. The positions of the pre-mRNA, the lariat-intron and the lariat-intron exon 2 intermediate are indicated.

IV.1.4 Discussion

The experiments described above strongly suggest that the Ssy5 protein, which was found as an A1 prey in a Slu7p two-hybrid screen, is not required for pre-mRNA splicing. Deletion of the *SSY5* ORF from the genome leads to a slow growth phenotype that becomes more pronounced at higher temperatures and leads to a severe temperature sensitivity at 37°C. For many splicing factors it has been shown that mutations in (or deletions of) genes encoding these proteins lead to temperature-sensitive growth phenotypes of the cells. In northern blot analyses of RNAs isolated from such mutant cells an accumulation of pre-mRNA is normally evident already two hours after shift to the non-permissive temperature (e.g. Maddock *et al.*, 1996, Vijayraghavan *et al.*, 1989). For the *SSY5* deleted strain YMA156/2, no pre-mRNA was detected after 5 hours incubation at 37°C, suggesting that the pre-mRNA was efficiently spliced even under non-permissive conditions. It has to be noted, that the YMA156/2 strain does not grow at all in rich liquid medium at 37°C (data not shown). From the *in vivo* data it could not be entirely ruled out, that Ssy5p might be a splicing factor required only for the second transesterification reaction to occur. If this was the case, one might not see an accumulation of pre-mRNA, because it would be efficiently processed into the intermediates, which are not detected in the northern analysis, since a probe for exon 1 was used. Due to the obvious loading differences of the samples a definite statement about the effect on the mRNA levels was also difficult to make. However, the result of the *in vitro* splicing assay strongly argues against this interpretation. Splicing intermediates do not accumulate in the absence of Ssy5p, and the products of the splicing reaction are produced to wt levels.

During the course of the experiments described above, some more information on the function of the Ssy5 protein became available from other work: after the initial identification of the *ssy5* gene in a screen for mutants sensitive to the sulfonylurea-herbicide MM, an inhibitor of the biosynthesis of branched-chain amino acids, the gene was cloned via complementation from a low copy number library (Joergensen *et al.*, 1998). In more directed experiments it was shown that the *ssy5* mutant strain was

deficient in the uptake of branched-chain amino acids such as isoleucine and valine, but that in addition the uptake of several other neutral amino acids, especially leucine and phenylalanine, was effected. If toxic analogs of these amino acids were added to the growth medium, the mutant strain showed resistance to these molecules, which gives further evidence, that amino acid uptake is reduced.

The structural gene for the branched-chain amino acids permease had been identified by Grauslund *et al.* (1995) as *BAP2*. Truncation of Bap2p at its carboxy-terminus leads to an **in**creased uptake of branched-chain amino acids. Joergensen and collaborators found, that the *ssy5* mutation is epistatic over the dominant Bap2p truncation, and they propose that Ssy5p might be involved in regulating the expression of *BAP2*.

If the Ssy5 protein is not involved in the splicing pathway, the question remains why it was found as interactor with two known splicing factors in the two-hybrid screens? The Ssy5 prey fusion proteins identified with the Slu7p- and the Isy1p bait do overlap to a large extend and they share a common region consisting of the amino acid residues 129-238. The region contains 33% charged residues, which might confer some nucleic acid binding capacity. Maybe this gives rise to an aspecific interaction with an RNA molecule that in turn interacts with Slu7p and/or Isy1p. If one thinks of a role of the Ssy5 protein as a regulator of *BAP2* expression, a nucleic acid binding activity could very well be imagined.

IV.2 Characterisation of the Yor275 protein

The Yor275 protein was found as an A3 interactor in each of two two-hybrid screens with the Slu7 protein as bait. Two distinct fusions were identified in these screens (see also figure III.4, section III.2.2). In addition, the protein was reported to represent an A1 prey candidate in a two-hybrid screen performed with the Snp1p bait, the yeast homologue of the human U1 snRNP-associated 70K protein (Fromont-Racine *et al.*, 1997 and Smith and Barrell, 1991).

No function had hitherto been assigned to the Yor275 protein, but it shows homologies to other proteins in yeast and other organisms, some of which have been functionally linked to signal transduction pathways. The conservation of the protein in different organisms will become important later and will be discussed in the final part of this chapter.

The two-hybrid interaction with two known splicing factors made this protein an interesting candidate for further investigation. In order to find out more about a putative role in the pre-mRNA splicing pathway, it was decided to use the Yor275p as a bait in a "second round" two-hybrid screen. Confirming the interactions with the Slu7p and/or the Snp1p in the reciprocal orientation or finding other interactions with components of the splicing machinery would give some evidence for an involvement of the Yor275 protein in the splicing process.

IV.2.1 Construction of the Yor275p bait

The *YOR275c* ORF was cloned into the pASΔΔBg vector (see table II.1.7), which should allow expression of a GBD::Yor275p fusion protein. The bait was constructed via an *in vivo* gap-repair cloning strategy, which has the advantage that only part of the reading frame has to be generated by PCR, therefore PCR-generated mistakes in the ORF sequence can be minimised. An overview of the cloning procedure is given in figure IV.5.

Figure IV.5:
In vivo **gap-repair cloning of bait plasmid pMA275.** 1) Oligonucleotide primers 275A and 275B were used to amplify and isolate an approximately 400 bp long fragment (275N) from genomic DNA encoding the N-terminus of Yor275p. 2) The fragment was then cut *Ecl*136II (E) and *Sal*I (S) and ligated into plasmid pASΔΔBg, previously linearised by *Sma*I (Sm), *Sal*I digestion, to produce plasmid pMA275N. 3) A DNA fragment containing the 300 3'-terminal base pairs (C) of ORF *YOR275c* and about 500 bp downstream of the ORF was isolated from plasmid p7-IV-75 by *Sal*I and *Nsi*I (N) restriction. 4) The fragment was then ligated into pMA275N, previously cut with *Sal*I and *Pst*I (P) producing plasmid pMA275NC. 5) pMA275NC was linearised by *Sal*I digestion, dephosphorylated and transformed into yeast strain BMA64, where it is repaired via homologous recombination to produce bait vector pMA275 6) pMA275 is isolated from yeast (plasmid rescue procedure) and its identity checked by restriction digests and DNA sequencing.
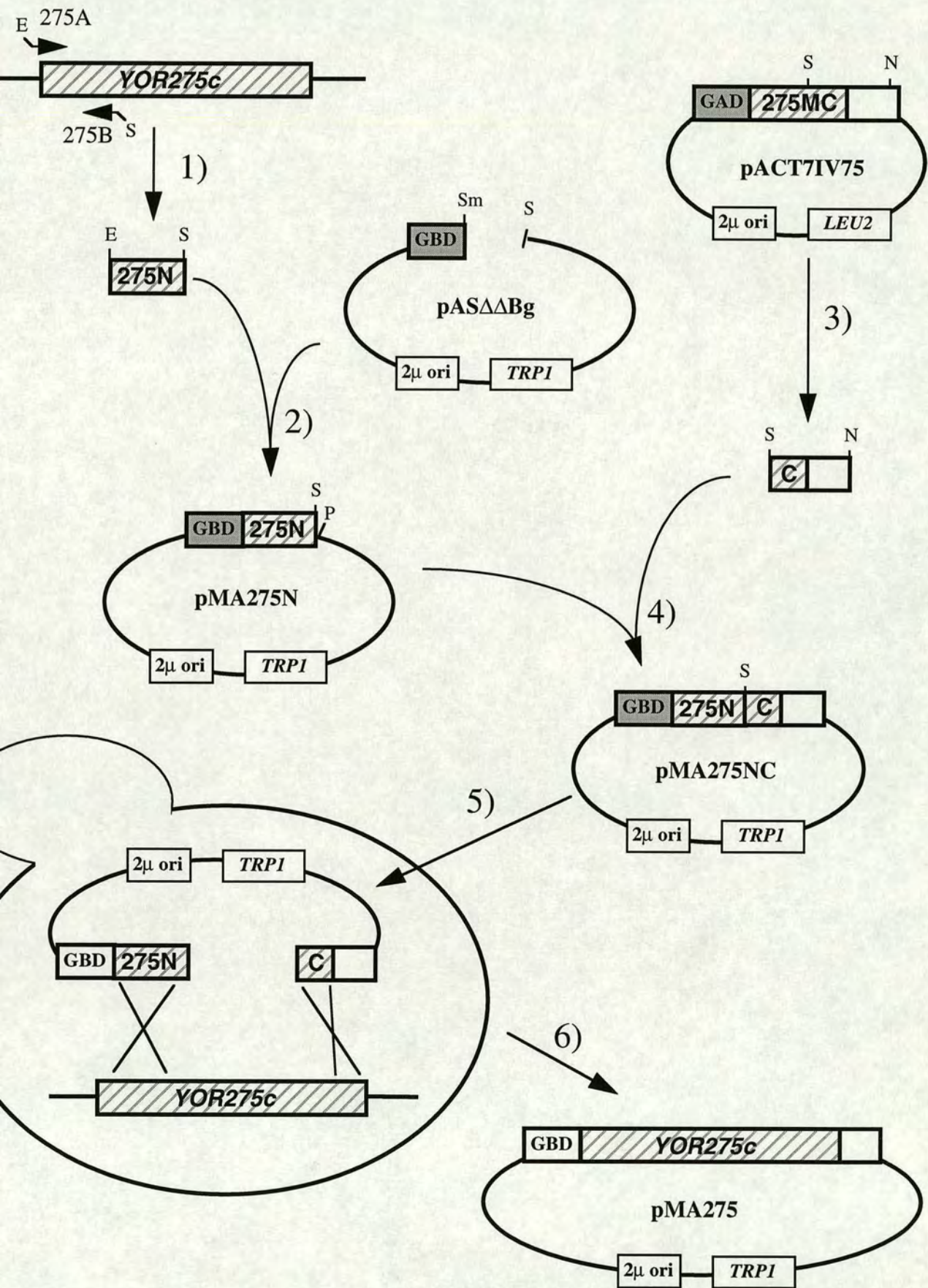
After the gap-repair in yeast strain BMA64, the plasmids were reisolated and transformed into *E. coli.*. Plasmid "Mini" DNA-preparations were made and the plasmids checked for a repair event by restriction digest. One clone out of 64 could be identified, in which, as determined from the restriction digest pattern, the *YOR275c* ORF had been successfully repaired. The plasmid was sequenced using the oligonucleotides N3027 (pAS2ΔΔ forward) and P5148 (pAS2ΔΔ reverse), in order to verify the repair event. This confirmed that the gap in the *YOR275c* ORF had been repaired and no PCR-generated errors had been introduced into the n-terminus. Subsequently, the plasmid, designated as pMA275, was transformed into yeast strain CG1945, in order to perform a two-hybrid screen using GBD::Yor275p as a bait.

IV.2.2 Small scale Yor275p two-hybrid screen

In order to check the bait construct for potential autoactivation before a full two-hybrid screen was performed, a small scale two-hybrid screen was set up using the EDYL. A mating efficiency of 8 % allowed the screening of 90000 diploid cells for the interaction of the GBD::Yor275p bait with prey fusions. After 3 days only a background of very small colonies was visible on the YMM -LWH plates. After 4 days 28 colonies were grown above background, picked and streaked onto fresh YMM -LWH medium. Two days later a β-gal overlay assay was performed on these cells, but none of them developed a blue colour even after 24 h of incubation. Thus, the bait fusion did not autoactivate transcription of the reporter genes and it seemed reasonable to use the bait in a full scale screen.

IV.2.3 The Yor275p two-hybrid screen

A full scale two-hybrid screen with the GBD::Yor275p-bait was performed using 1 vial of the FRYL. A total of 72 million diploid clones were screened. The mating efficiency was calculated as 33%. After 3 days the plates were overlaid with an X-gal agar mix and 37 blue colonies were picked after 3 h (clones 1-37), another 11 after 7 h

(clones 38-49) and an additional 8 light blue colonies after 19 h (clones 50-57). The positives were restreaked onto YMM -LWH plates and the interaction was assayed again, this time in a β-gal filterlift assay. The relative strength of the interactions (intensity of the blue colour in the filterlift assay) was recorded and is listed in table IV.1, which summarises the results of this two-hybrid screen. Clones that do not appear in the table did not regrow after the overlay assay.

The GBD::Yor275p bait behaved very specificly in the two-hybrid screen. Only four prey proteins were identified that fall into the A category. The candidates Ybr216p and Ydr332p are hypothetical proteins, which were each found only once and must be classified as A3/A2 and A4 preys, respectively. These proteins do not share significant homologies to other proteins in the databases.

The other two prey proteins both fall into the A1 category. Mrpl19p, encoded by the ORF *YNL185c* was identified 20 times in 5 independent fusions. Rim1p, encoded by ORF *YHL027w*, was found 14 times in 6 independent fusions. Figure IV.6 shows the fragments of Mrpl19p and Rim1p which interacted with Yor275p in the two-hybrid screen.

The carboxy-terminal 29 amino acid residues of the Mrpl19 protein are sufficient to support the two-hybrid interaction with the Yor275 protein. Mrpl19p (also designated YmL19) is a 16 kDa protein that is part of the large subunit of mitochondrial ribosomes and shares extensive homologies to prokaryotic ribosomal proteins of the L11 subfamily (see Graack and Wittmann-Liebold, 1998 and references therein).

For the Rim1 protein, the c-terminal third of the protein (amino acids 404-625) is common to all prey fragments identified, and thus, is likely to represent the region of interaction with the Yor275 protein. The Rim1 protein has been shown to be a transcriptional regulator of the *IME1* gene, which encodes a protein that activates the expression of a number of early meiotic genes (Su and Mitchell, 1993). The activity of Rim1p requires proteolytic cleavage at the C-terminus and this cleavage is stimulated under alkaline growth conditions. Three proteins, Rim8p, Rim9p and Rim13p are thought to modulate Rim1p function, probably via a signal transduction cascade (Li and Mitchell, 1997).

Table IV.1:
Results of the Yor275p two-hybrid screen. A Gal4 DNA-binding domain-Yor275 fusion protein (GBD::Yor275p) was used as bait to screen the FRYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. No. - frequency with which a fragment was identified in the screen; Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - approximate insert length (nucleotides) determined by restriction digest of the prey plasmid with *Bam*HI and subsequent agarose electrophoresis; preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size (bp)	Insert size (bp)	Category	Protein info
45	1		YBR216c	2	c	82	2024	1300	A3/A2	
12	1		YDR332w	4	w	838	2069	400	A4	
57	1	RIM1	YHL027w	8	w	1015	1877	1200	A1	Meiotic regulator
17	5	RIM1	YHL027w	8	w	1039	1877	800	A1	Meiotic regulator
55	1	RIM1	YHL027w	8	w	1069	1877	800	A1	Meiotic regulator
3	4	RIM1	YHL027w	8	w	1159	1877	700	A1	Meiotic regulator
6	2	RIM1	YHL027w	8	w	1192	1877	700	A1	Meiotic regulator
20	1	RIM1	YHL027w	8	w	1210	1877	1100	A1	Meiotic regulator
47	1	MRPL19	YNL185c	14	c	112	476	500	A1	Mitoch. ribosomal protein
10	8	MRPL19	YNL185c	14	c	238	476	500	A1	Mitoch. ribosomal protein
7	4	MRPL19	YNL185c	14	c	304	476	400	A1	Mitoch. ribosomal protein
19	3	MRPL19	YNL185c	14	c	307	476	800	A1	Mitoch. ribosomal protein
11	4	MRPL19	YNL185c	14	c	324	476	500	A1	Mitoch. ribosomal protein
2	1	TY2-1							B	
4	1	TY2-1							B	
29	1	TY2-1							B	
31	1	TY2-1							B	
22	1	TY1-2							B	
27	1	TY1-2							B	
49	1		e.g YPR204w						B	Subtelomeric encoded proteins, ATP/GTP-binding ?

Figure IV.6:
A1 category preys identified in the Yor275p two-hybrid screen.
The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen. **A)** Mrpl19p prey fragments. **B)** Rim1p prey fragments..

A)

clone: fragment: No.:

47:     **Mrpl19p(38-158)**   1 x

10:     **Mrpl19p(80-158)**   8 x

7:      **Mrpl19p(102-158)**   4 x

19:     **Mrpl19p(103-158)**   3 x

11:     **Mrpl19p(129-158)**   4 x

        Mrpl19p

        1   129   158

B)

clone: fragment: No.:

57:     **Rim1p(339-625)**   1 x

17:     **Rim1p(347-625)**   5 x

55:     **Rim1p(358-625)**   1 x

3:      **Rim1p(387-625)**   4 x

6:      **Rim1p(398-625)**   2 x

20:     **Rim1p(404-625)**   1 x

        Rim1p

        1   404   625
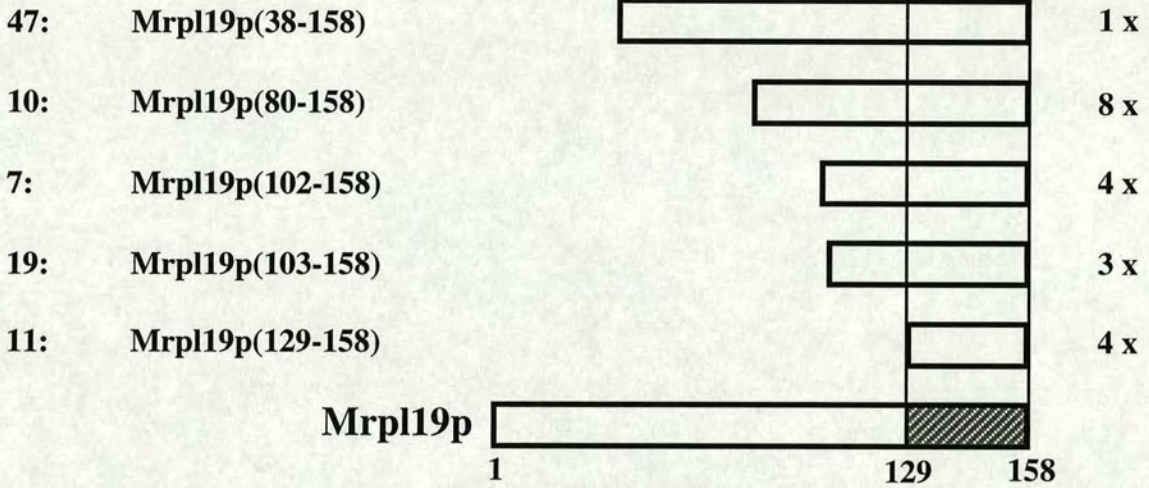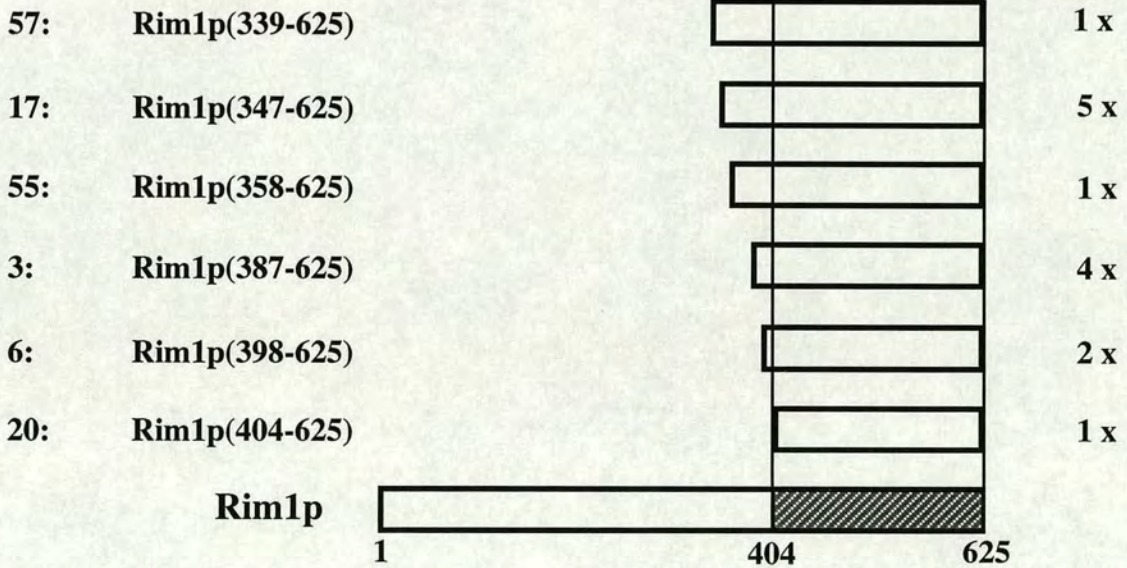
Database homology searches give some evidence that the observed Yor275p-Rim1p two-hybrid interaction is indeed of functional relevance:

The Rim1 protein shares 22% amino acid identity over the whole protein sequence with the PacC protein of *Aspergillus nidulans* (*A. nidulans*). In particular, a region containing three C2H2 Zn-finger motifs is highly conserved between the two proteins (with 64% identity and 86% similarity among 86 amino acids). Figure IV.7 shows a sequence alignment of Rim1p and the PacC protein.

The two proteins also seem to be functionally related: at alkaline ambient pH, PacC activates transcription of alkaline-expressed genes (including itself) and represses transcription of acidic-expressed genes (Tilburn *et al.*, 1995). Furthermore, as for the Rim1p, the activation of the PacC protein in response to an alkaline pH requires proteolysis at the c-terminus (Orejas *et al.*, 1995). In addition, proteins have been identified (PalA, B, C, F, H and I) that might act as upstream regulators of PacC in a signal transduction cascade as do Rim8, 9 and 13p for Rim1p (Shah *et al.*, 1991; Negrete-Urtasun *et al.*, 1997 and Arst *et al.*, 1994). The Rim9p is homologous to the PalI protein and both proteins share four hydrophobic, putative membrane spanning domains. Therefore, these proteins might represent the most upstream elements of the signal transduction cascade, that sense the pH outside the cell (Denison *et al.*, 1998). Recently, it was reported that the Rim13 protein actually is a calpain-like cysteine protease and it was renamed Cpl1p (Futai *et al.*, 1999). The authors show that disruption of the *CPL1* gene leads to Rim1p degradation and suggest that it might function in stabilising Rim1p. Cpl1p is homologous to the PalB protein of *A. nidulans*.

Figure IV.7:
Sequence alignment of Rim1p with the *Aspergillus nidulans* **PacC protein.** Sequences were aligned using the PILEUP program of the GCG9 suite of sequence analysis programs, with identities and similarities highlighted using BOXSHADE 3.21 (section II.4). Black boxes highlight identical residues, and grey boxes indicate the conservation of the nature of the amino acid at that site. Accepted conservative groupings were M=I=L=V, K=R=H, F=Y=W, S=T, E=D, A=G and Q=N.

```
PacC      1  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~MSEPQDTTTAP.STTAAPMP.TSTSQ.DS
Rim1p     1  MVPLEDLLNKENGTAAPQHSRESIVENGTDVSNVTKKDGLPSPNLSKRSSDCSKRPRIRCTTEAIGLNGQEDERMSPGSTSSSCLPYHSTSHLNT

PacC     27  P.......SAQQPAQVSSATAASAAATAAAASAAVANPPMNGTTTRPSEELSCLWQGCSEKCPSPEALYEHVCERHVGRKSTNNLNLTCQWGSCR
Rim1p    96  PPYDLLGASAVSPTTSSSSDSSSSSPLAQAHNPAGDDDDADNGDSEDITLYCKWDNCGMIFNQPELLYNHLCHDHVGRKSHKNLQLNCHWGDCT

PacC    115  TTTVKRDHITSHIRVHVPLKPHKCDFCGKAFKRPQDLKKHVKTHADDSVLV...RSPEPGARNPDMMFGGGAKGYATAAHYFEPALNAVPSQGYA
Rim1p   191  TKTEKRDHITSHLRVHVPLKPFGCSTCSKKFKRPQDLKKHLKIHLESGGILKRKRGPKWGSKRTSKKNKSCASDAVSSCSASVP..SAIAGSFKS

PacC    207  HG.APQYYQSHPPPQPANPSYGNVYYALNHGPEAGHASYESKKRGYDALNEFFGDLKRRQFDPNSYAAVGQRLLGLQSLSLPV.LSSGPLPEYQP
Rim1p   284  HSTSPQIL....PPLPVGISQ............HLPSQQQQRAI.SLNQLCSD.ELSQYKPVYSPQLSAR...LQTILPPLYYNNGSTVSQGA

PacC    300  MPAPVAVGGGGYSPGGAPSAPAYHLPPMSNVRTKNDLINIDQFLQQMQDTIYENDDNVAAAGVAQP..GAHY..VHGGMSYRTTHSPPTQLEPSH
Rim1p   356  NSRSMNVYEDGCSNKTIANATQF.FTKLSRNMTNNYI..LQQSGGSTESSSSSGRIPVAQTSYVQPPNAPSYQSVQGGSSISATANTATYVEVRL

PacC    391  ATATSSASMMPN.PATHSPSTGTPALTPPSSAQ.SYTSGR.SPVSLPSATRVSPPHHEGGSMYPRLPS.ATMADSMAAGYPTASSTAPPSTLGGI
Rim1p   448  AKYPTGPSLTEHLPPLHSNTAGGVFNRQSQYAMPHYPSVRAAPSYSSSGCSILPPLQ...SKIPMLPSRRTMAGGTSLKPNWEFSLNQKSCTNDI

PacC    482  F.......DHDDRRRYTGGTLQRARPETRQLSEEMDLTQDSKDEGERTPKAKEHSSPSSPERISASLIDPALSGTAAEAEATLRTAQAATEVAER
Rim1p   540  IMSKLAIEEVDDESEIEDDFVEMLGIVNIIKDYLLCCVMEDLDDEESEDKDEENAF......LQESLEKLSLQNQMGTNSVRILTKYPKILV~~~

PacC    570  ADVQWVEKVRLIEYLRNYIASRLERGEFENNESGGGNSSSNGSSHEQTPEASPDTHMEGVESEVPSKAEEPAVKPEAGDVVMYPTLRAVDEDGDS
Rim1p   626  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

PacC    665  KMP
Rim1p   626  ~~~
```

Negrete-Urtasun *et al.* (1997) noted a homology of the PalA protein with the Yor275 protein of *S. cerevisiae*. A GeneStream alignment shows a 25% amino acid identity of the proteins over the whole sequence. In particular the n-terminal halves are very well conserved (37% identity and 64% similarity among 325 amino acids). Figure IV.8 shows a sequence alignment of the Yor275p and the PalA protein.

The finding that Yor275p interacts with the Rim1 protein in the two-hybrid screen, is the first evidence that it could indeed be functionally related to the PalA protein and that the Yor275p-Rim1p interaction in *S. cerevisiae* resembles the PalA-PacC interaction in *A. nidulans*. Furthermore, it supports the hypothesis that the two conserved signal transduction pathways might have evolved from a common ancestor.

Figure IV.8:
Sequence alignment of Yor275p with the *Aspergillus nidulans* PalA protein. Sequences were aligned using the PILEUP program of the GCG9 suite of sequence analysis programs, with identities and similarities highlighted using BOXSHADE 3.21 (section II.4). Black boxes highlight identical residues, and grey boxes indicate the conservation of the nature of the amino acid at that site. Accepted conservative groupings were M=I=L=V, K=R=H, F=Y=W, S=T, E=D, A=G and Q=N.

```
Yor275p    1  ~MSELLAIPLKRTLEVDFATELSKLIDTTSFQTASFFQSDILKVVDARNNAI.APDISIDGLSALKEYYVILLQLEKKFP.NNQIEFTWFQTLS.
PalA       1  MASNILQIPFRRSHTVSLSTALTQYISTKYDQRPDMFADDLLIIDRLRNEAINVQEPHVSGISRLVTYAAQLKWLGGKFPVDVGVEFPWYPAFGF

Yor275p   92  QKSRGTSQYSLQWEKLTIIYNIGCMYSLLALNSNNDAAESLKTSCLYFQNAAGCFKHV.LDHQKNLETIPV..VDDATLNALTSLMLAQAQECFW
PalA      96  NTSRPVSQDNIRFELANVIFNLAALYSQLAFAVNRTTTDGLKQACNYFCQAAGILAHLRTDIVPDMRSAFPEDMDEMTLRSLEELLLAQAQECFW

Yor275p  184  FKAVQDKHKDSLIAKLSQQIVDFYCEAINDAQRGKLIRSDWINHLKAKKAYFSAVTYYRIALSFNEKKQFGNVVKALQMGLQFINESTLSSQAKF
PalA     191  QKAVMDGLKDASIARLAGQVSDFYGDACDHAVKSNAISPEWIHHMTAKQHHFAAAAQYRQSLDCLEKRKYGEEVARLRDAVACVNEALKESRWIN

Yor275p  279  KTV..........VESSLKEAQRDNEFIYLQEVP..SELPSIKPALMVKPSSSATLLPSIKK.......DETLFKDLIPIEVMEYCTAYNERQDE
PalA     286  RTVLGDLQGLKNRVTEDLKRAEKDNDMIYLNPVPPKSELKLIDRACMVAAKAPSQVTDAISMLGEKGPLGQPLFSKLVPYAVHIAASIYSDRRDR

Yor275p  355  YVEQRVTNPLASLNKLLKESLTTFQIP...QGLTK.........VSEAELSHYQASLNNL...LINNKNVQ.....VQLDNIEQILNEEAFTDNQL
PalA     381  LVNERIIGELENMTDKLRDLLSSLNLPGSLQALEKPLGLPPSLVAHAEEMRQQDGLNRLRKSLLDIAKVKSNDRAVYTEGVELLAAEKA.EDDAS

Yor275p  431  RLKHGTLNWTLPESSTTNTAYYEKLKKLRGYLDEGSAIDKQTNELF...QSIDKNLIGS....EIRLPESN........DPLTNKIKMIIQERND
PalA     475  RRKFGTDRWTREASEAAAPKLYTTAREIDGYFTSAQSSDNLVEQKLHDSEAVFRVLTGTNRDLEAFVPSSRRATIPPEVEREVSRLRSCISEVNR

Yor275p  511  YIDRTRRK.............SSEYRILPKIITSYKKNG..TVDFEPIFIGHLKYFDEDLRYVNSTKEENIKLIEEV.NLSKKNNPGRSGIEPK
PalA     570  LESRRKRKAQAVKDKARADDISSALVREAARLEREFPMQAIQASQFEDLFESRLRDYDVDLDMVAQEMHDQDQIVAQVRDANRAFTRAHTGDAST

Yor275p  589  KMVRIDPRELYIEDLRYSFKLLDEVKENLSAGTAFYENLITSTSNLYNEVQEYDTARRAEKARLDKSLTFEDQ~~~~~~~~~~~~~~~~~~~~~~~~
PalA     665  K.....EREKALQELENGYLKYKEIISNIEVGRKFYNDLAKIVGRFRDDVKAFVHKRRMEASQLEQDISSVAAMASLNISPIRQPPQQTVVSAPV

Yor275p  662  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
PalA     755  SVSAAASVPAPTHFNPVKPQPQPPSQAIPPQSQPQGEPQPCGHL
```

IV.2.4 Discussion

The Yor275p was found to interact in a two-hybrid screen with the Slu7p bait. In addition it was identified with another well characterised splicing factor, Snp1p, in a two-hybrid screen performed in Pierre Legrain's laboratory (Fromont-Racine *et al.*, 1997). On one hand it seems promising, if a protein is found in multiple two-hybrid screens with known splicing factors, this might suggest a contact with the spliceosome and an involvement in the splicing pathway. On the other hand, if the baits that were used in these screens, are known to act at different stages of the splicing reaction and no direct physical link has been established between these stages, it could be a concern that the observed two-hybrid interactions are based on non-specific contacts without functional relevance. In case of the Slu7p-Yor275p and the Snp1-Yor275p interactions, we have such a situation: the Snp1p, the yeast homologue of the human U1-70K protein, is tightly associated with the U1 snRNP but not with other snRNPs and is required for commitment complex formation (Smith and Barrell, 1991; Neubauer *et al.*, 1997). The Slu7p, on the other hand, is known to have a function during the second step of the splicing process (e.g. Frank and Guthrie, 1992; Umen and Guthrie, 1995a) and no data are available to date that suggest a function earlier on. In fact, immunoprecipitation experiments with anti-Slu7p antibodies strongly suggest that the protein joins the spliceosome after the action of Prp16p, i.e. just before the second transesterification reaction (Brys and Schwer, 1996). Thus it can be expected, that Snp1p has left the spliceosome together with the U1 snRNP, before the Slu7p associates with it. A complex between the three proteins Snp1p-Yor275p-Slu7p is therefore unlikely to be found in the spliceosome. However, it could be that Yor275p interacts at different times with Slu7p and Snp1p.

The results of the two-hybrid screen argue against such an interpretation, in fact they suggest that Yor275p is not involved in pre-mRNA splicing at all. Neither interaction with Slu7p nor Snp1p was found in the reciprocal orientation, nor any other interactions with known splicing proteins. The two-hybrid screen with the Yor275p bait gives the first experimental evidence for an involvement of the protein in a signal transduction pathway that is dependent upon ambient pH and that regulates sporulation in budding yeast. Two findings support this assumption: first, Yor275p interacts

significantly with Rim1p, a transcriptional activator of the meiotic regulatory gene *IME1*, which is the target gene at the end of the signal transduction cascade. Second, this pathway seems to be conserved and can be found in *A. nidulans*, where it serves to regulate the expression of genes that allow the fungus to respond to different ambient pH conditions. The Rim1p, Rim9p, Cpl1p and the Yor275p of *S. cerevisiae* and the PacC, PalI, PalB and the PalA proteins of *A.nidulans*, respectively are homologous and could well perform similar roles in these pathways. Figure IV.9 gives an overview of the two conserved pathways.

The strong two-hybrid interaction between the Yor275p and the Rim1p supports the idea, although it does not prove it, that these proteins directly contact each other. The interacting region, common to all prey fragments found, comprises the amino acid residues 404-625. Thus, Yor275p might bind to the c-terminal fragment of Rim1p (amino acids 531-625) that is cleaved off for the activation of the protein. Although there is no reason to believe from the primary structure of Yor275p that it itself is the protease that activates Rim1p, it could be well imagined that the protein is directly involved in the proteolytic step.

Given the evidence presented above for the function of the Yor275p in the regulation of meiosis and sporulation, the second A1 interaction found in the two-hybrid screen, namely the interaction with the mitochondrial ribosomal protein Mrpl19p, is likely to be non-physiological. Why Yor275p interacts so strongly with a small c-terminal region of the Mrpl19 protein cannot be easily predicted, since this region does not contain any recognisable motifs or domains.
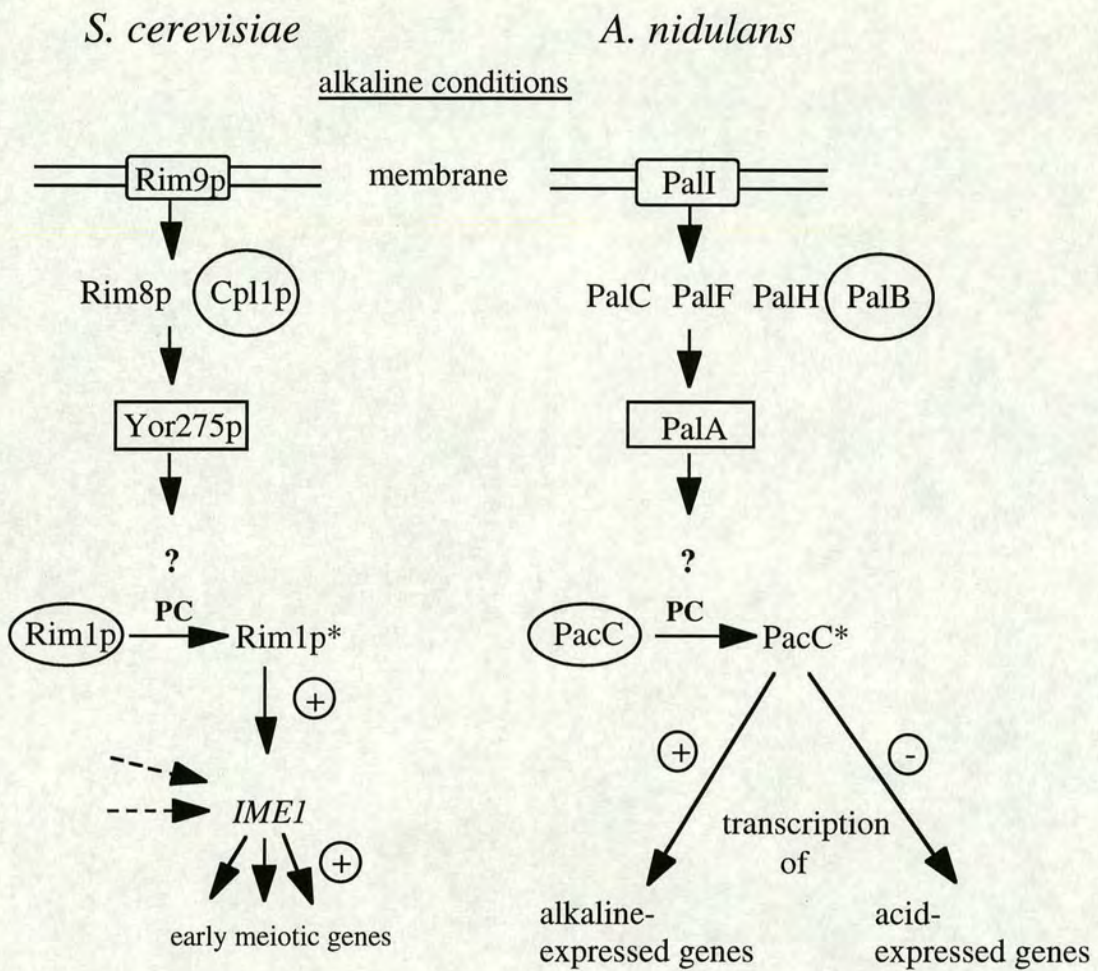
Figure IV.9:
Putative order of events in two related signal-transduction pathways in
S. cerevisiae and Aspergillus nidulans under alkaline conditions.
Proteins with significant sequence similarity are boxed in the same shape; +
transcriptional activation; - transcriptional repression; PC - proteolytic cleavage (? the
responsible protease has not yet been identified); * - active protein; dashed arrows
indicate other signal transduction cascades which are believed to influence *IME1*
transcription; (see text for discussion).

IV.3 Characterisation of the Ymr44 protein

The 55 kDa Ymr44 protein was identified in a two-hybrid screen with the Prp18p bait as an A4 candidate (Screen M18, table III.3). The protein was chosen for further analyses for two reasons: first, in a previous Prp18p two-hybrid screen performed in our laboratory, a different fragment of the protein was found, that shared an overlapping region with the fragment found in screen M18 (R. van Nues, personal communication). Second, the protein was identified in two-hybrid screens with three other splicing factors within the TAPIR network (Fromont-Racine *et al.*, 1997 and Pierre Legrain, personal communication). The significance of these interactions will be discussed later in this chapter.

To get first clues about the function of the protein, a gene deletion was performed to find out, whether the protein is required for cell viability. If not, the deletion might cause a temperature-sensitive phenotype, that could then be further investigated for a defect in the splicing process under non-permissive conditions. Furthermore, the protein was used as bait in a two-hybrid screen in order to see whether the observed two-hybrid interactions with known splicing proteins could be reciprocated or whether other interactions could be found that might suggest a function of Ymr44p unrelated to pre-mRNA splicing.

IV.3.1 Deletion of the *YMR44w* open reading frame from the genome

In order to delete the *YMR44w* ORF from the genome, oligonucleotides 44ko1 and 44ko2 (see table II.1.6) were used in a PCR to amplify the *HIS3* marker gene from plasmid YIp1. The 1.15 kb long PCR-product was gelpurified and transformed into the diploid strain BMA38, which is deleted for the *HIS3* ORF. The 3' ends of the oligonucleotides were chosen homologous to sequences just upstream and downstream of the *YMR44w* ORF, and thus should allow the PCR product to recombine into the *YMR44w* locus, replacing the ORF with the *HIS3* marker. Histidine prototrophic transformants were investigated by yeast colony-PCR for correct integration of the replacement-cassette into the *YMR44w* locus (see figure IV.10) By using

oligonucleotides 44ko1 (with the 3'-end homologous to the *HIS3* locus) and 44B (homologous to a region of the complementary strand just downstream of the *YMR44w* ORF), it could be confirmed that, in 7 out of 9 transformants which were tested, a replacement of the *YMR44w* ORF by the *HIS3* marker gene had taken place. Subsequently, the PCR-products were purified using the QIAquick PCR-purification kit (Qiagen) and sequenced using oligonucleotide 44B. The identity of the PCR product was confirmed.
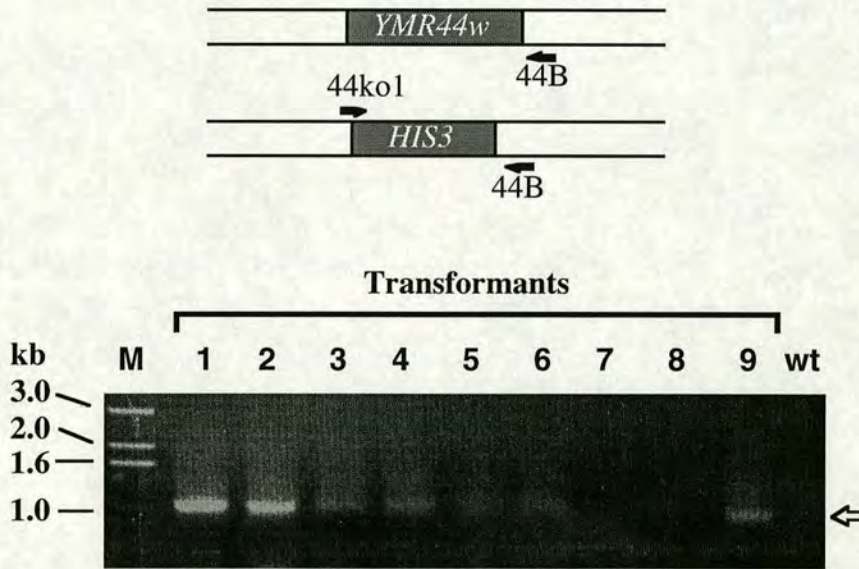
Figure IV.10:
PCR on yeast transformants to test for integration of the *HIS3* marker gene into the chromosomal *YMR44w* locus. After transformation of the BMA38 strain with a linear PCR product containing the *HIS3* marker gene flanked by approximately 40 base pairs of the YMR44w locus at either site, histidine prototrophs were streaked out onto fresh YMM -H medium for colony purification. Then a single large colony of the transformants and of the wt parental strain BMA38 was suspended in 0.02N NaOH, boiled for 5 minutes and an aliquot used in a PCR (II.3.2.8.2) using oligonucleotide primers 44ko1 and 44B. The positions on the template at which the primers anneal are indicated in the upper schematic drawing of the*YMR44w*-locus (either wild-type or after *HIS3*-integration). Aliquots of the PCR reactions were run on a 1.0 % (w/v) agarose gel which has been photodocumented. (The white arrow indicates the position of the amplified 1.15 kb fragment, which would be expected following successful integration of the *HIS3* cassette into the YMR44w target locus.)

A diploid transformant that had proven positive for the gene replacement (designated strain YMA44/1) was then sporulated. The generated asci were dissected onto YPDA agar plates and the separated spores incubated at $23^{\circ}C$ for 3 days. The result of the tetrad dissection is depicted in figure IV.11 A.

Out of 11 tetrads dissected, in 5 cases all four spores grew to colonies of comparable size, in four cases only 3 spores germinated and in two cases, two spores grew to colonies. The fact, that for 9 of 11 tetrads, more than 2 spores formed colonies suggested strongly that the *YMR44w* gene is not essential for cell viability. Furthermore, since for the majority of the spores grew to similar size, the gene does not seem to be required for cell growth.

To test this further, the spores of the tetrads, in which all four spores germinated were streaked onto histidine-lacking plates. In all cases, as expected only 2 spores grew (data not shown). Cells from 8 of these histidine prototrophic haploids (deleted for *YMR44w*) were serially diluted into microtiter plates and spotted onto YPDA agar plates using a pronged metal inoculator. The plates were incubated at a range of temperatures for 2 days (figure IV.11 B). This confirmed, what was already indicated by the tetrad analysis, namely that the deletion of the *YMR44w* gene from the genome, does not have a visible effect on cell growth in rich medium. The *ymr44w* deleted strain, designated YMA44/2, grows at a rate comparable to the parental wt strain at $23^{\circ}C$, $30^{\circ}C$ or $37^{\circ}C$.

A)

B)

23 °C   30 °C   37 °C
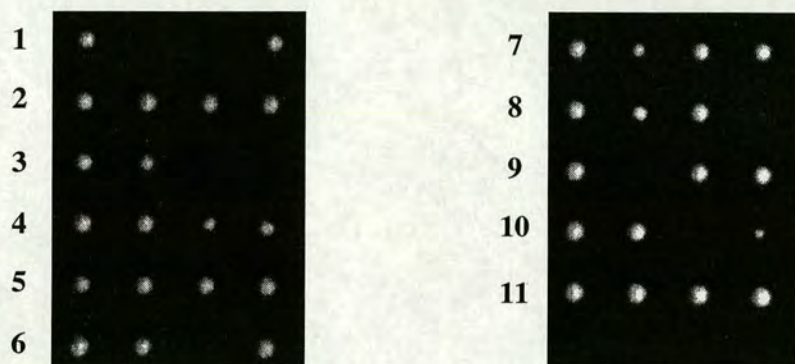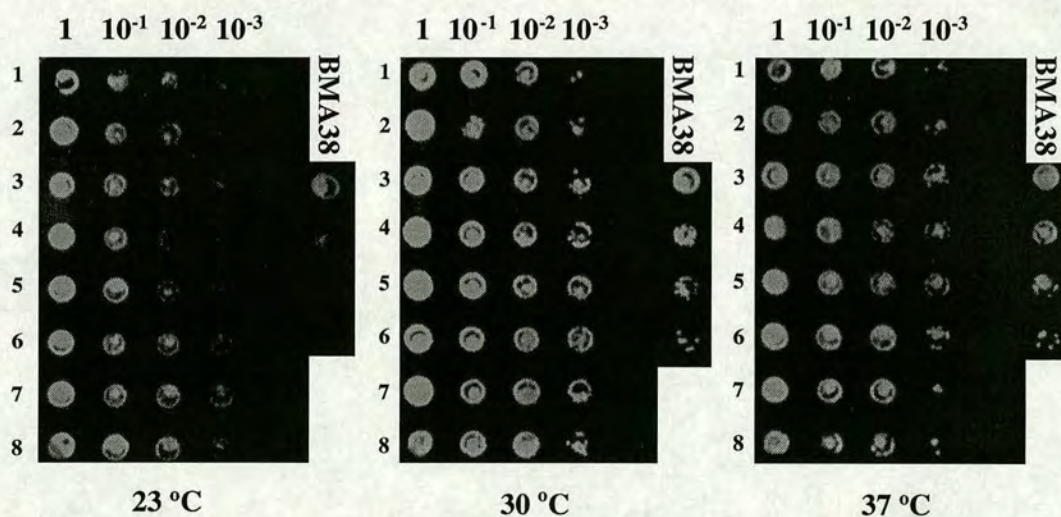
Figure IV.11:
Analyses of the effects of *YMR44w* deletion.
A) Tetrad analysis of strain YMA44/1. Strain YMA44/1 was grown overnight at 30°C in YPDA, the cells collected and transferred to Spm sporulation medium. After 5 days of incubation at 23°C the formed tetrads were dissected onto YPD agar plates. The spores were incubated for 3 days at 23°C.
B) Haploid cells deleted for the *YMR44w* locus (strain YMA44/2) as well as cells from the parental wild-type strain BMA38 were suspended in dilutions in microtiter plates, spotted onto YPDA agar plates and incubated at different temperatures for 2 days. (BMA38 cells were diluted in the same way as the cells from strain YMA44/2; this dilution series is depicted here in vertical orientation)

150

IV.3.2 Cloning of the *YMR44w* bait vector

In order to perform a two-hybrid screen with the Ymr44 protein, the *YMR44w* ORF had to be cloned into the bait vector. For this, the *YMR44w* ORF was amplified by PCR from genomic DNA of strain BMA38n using oligonucleotides 44A and 44B (table II.1.6). The 1.4 kb PCR product was then cut with *Ecl*136II and *Sal*I and ligated into vector pASΔΔBg, which had been cut with *Sma*I and *Sal*I in the multiple cloning site. The insert was sequenced twice over its entire length using oligonucleotides N3027 (pAS2ΔΔ forward) and P5148 (pAS2ΔΔ reverse) and no deviations to the wt gene sequence were detected. Also, by sequencing the junction between the GBD sequence and the start of the *YMR44w*-ORF, it was confirmed that the fusion was in the right frame to express the full-length fusion product. The plasmid was designated pMA44AS and was transformed into yeast strain CG1945 for use in the two-hybrid screen. However, the strain containing the bait plasmid flocculated badly in YMM -W liquid medium and was unsuitable for use in the screen. Thus, the cloning procedure was repeated, but this time the *Ecl*136II and *Sal*I cut PCR product was ligated into the pBTM116(-1) vector, which had been linearised before by digestion with *Sma*I and *Sal*I. Again integrity of the vector was checked by DNA-sequencing, for which the oligonucleotides W2248 (lexA forward) and the lexA reverse oligonucleotides were used. The resulting bait vector pMA44BTM was transformed into yeast strain L40. This time was no problem with flocculation of the cells when grown in YMM -W liquid medium. The LexA::Ymr44p (LA::Ymr44p) bait produced from this vector could thus be used to screen the two-hybrid library for interacting proteins.

IV.3.3 The Ymr44p two-hybrid screen

The Ymr44p bait was tested in a directed two-hybrid mating assay against 13 different prey fusions of known splicing factors, but no interactions could be detected (data not shown). This suggested that the bait did not (strongly) autoactivate and that it was suitable to be used in a full-scale two-hybrid screen.

The FRYL was screened for interactors with LA::Ymr44p. A mating efficiency of 64% allowed screening of 120 million diploid cells for transcription of the *HIS3* reporter gene. No 3AT was added to the selective plates in this screen. Table IV.2 summarises the results of the screen. Four proteins were identified as A1 prey in this two-hybrid screen. By far the strongest interactor, with respect to the number of fusions identified in the screen, is the 104 kDa Fir1 protein, which is encoded by ORF *YER032c* (figure IV.12). The common region shared by all 5 prey fragments is confined to a small carboxy-terminal peptide sequence, namely to the amino acid residues 766-832 of the 925 amino acid long protein. There are no obvious sequence motifs within this region, that could give an indication for a protein-protein or protein-RNA binding capacity of this fragment.

Fir1p was originally isolated by Russnak *et al.* (1996) in a two-hybrid screen using the polyadenylation factor Ref2p as bait (thus the name: **F**actor **i**nteracting with **R**ef2p). Ref2p was shown to be directly involved in mRNA 3'-end formation. *In vitro*, it is required for efficient use of weak poly(A) sites. Endonucleolytic cleavage occurs accurately, but at significantly lower rates in *ref2Δ* strains (Russnak *et al.*, 1995). Fir1p, like Ref2p, is not essential for cell viability, but deletion of the *FIR1* gene leads to temperature-sensitive growth and furthermore to a reduction of the usage of suboptimal poly(A) sites *in vivo* (Russnak *et al.*, 1996). In *Δfir1/Δref2* double mutants this effect is exacerbated, supporting the idea of an involvement of the two proteins at the same stage of 3'-end processing.

Table IV.2:
Results of the Ymr44p two-hybrid screen. A LexA-Ymr44 fusion protein (LA::Ymr44p) was used as bait to screen the FRYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. No. - frequency with which a fragment was identified in the screen; Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - insert length (nucleotides) determined by sequencing, * - insert size has been determined approximately by *Bam*HI digestion of the prey plasmid and subsequent agarose electrophoresis. Preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size (bp)	Insert size (bp)	Category	Protein info
44-1	5	HEX3	YDL013w	IV	w	13	1860	1200*	A3	Hexose transporter
44-5	1		YBR223c	II	c	158	3438	989	A1	
44-6	1		YBR223c	II	c	252	3438	1821	A1	
44-9	1	FIR1	YER032c	V	c	1807	2778	776	A1	Polyadenylation factor
44-11	1	FIR1	YER032c	V	c	1957	2778	539	A1	Polyadenylation factor
44-29	1	FIR1	YER032c	V	c	1988	2778	753	A1	Polyadenylation factor
44-14	1	FIR1	YER032c	V	c	2125	2778	441	A1	Polyadenylation factor
44-30	2	FIR1	YER032c	V	c	2296	2778	330	A1	Polyadenylation factor
44-32	1	SAP1	YER047c	V	c	442	2694	500*	A4	AAA-ATPase family
44-18	3		YHR134w	VIII	w	406	810	800*	A1	
44-19	1		YHR134w	VIII	w	268	810	600*	A1	
44-10	2		YNL078w	XIV	w	246	1224	1200*	A1	
44-20	1		YNL078w	XIV	w	988	1224	500*	A1	
44-24	1	anti	(YGR187c)		w			n.d	B	
44-23	1		YLL066c and others					n.d	B	Subtelomeric repeated prot.
44-27	2							n.d	B	Ty1-element

clone:	fragment:		No.
44-9:	Fir1p(603-861)		1 x
44-11:	Fir1p(653-832)		1 x
44-29:	Fir1p(664-913)		1 x
44-14:	Fir1p(709-855)		1 x
44-30:	Fir1p(766-875)		2 x
Fir1p		1 766 832 925	
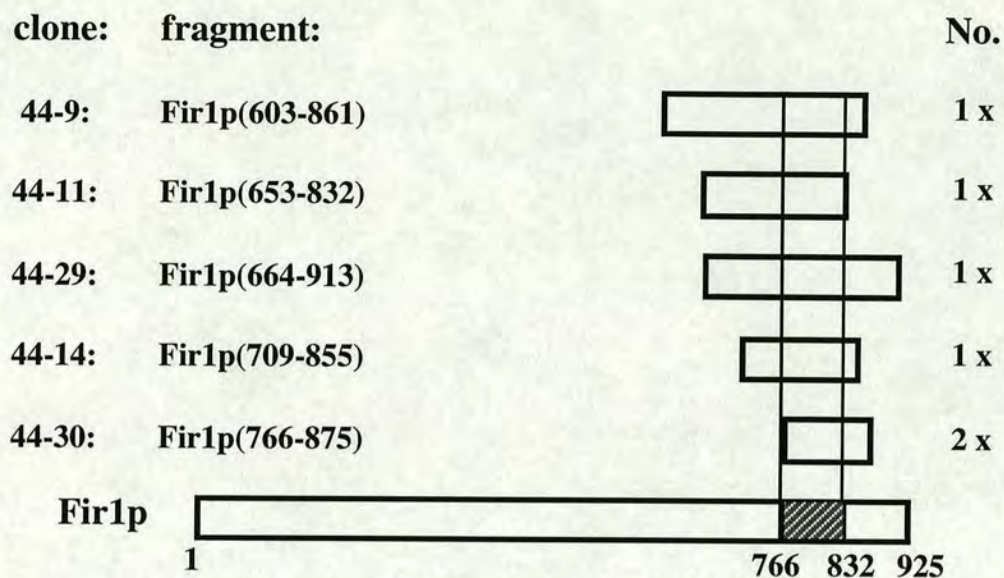
Figure IV.12:
Fir1p prey fragments identified in the Ymr44p two-hybrid screen. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets define the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen.

The other A1 prey candidates identified with the Ymr44w bait are hypothetical proteins, encoded by ORFs *YBR223c, YNL078w* and *YHR134w*, respectively (figure IV.13).

The hypothetical 62 kDa Ybr223 protein was found twice as independent fusions, which share a small region of overlap from amino acid 85 to approximately amino acid 120-150 (the 3' ends of the library plasmid inserts have not been sequenced in this case). This region contains no recognisable motifs and exhibits no significant homologies to other proteins in the databases. The amino acids 98-542 of the protein exhibit a 22% identity with a *C. elegans* protein of unknown function (Acc.no. AAC68960).

The Ybr223 protein was also identified as an A1 candidate in a two-hybrid screen with the Rpc53 protein, an essential subunit of RNA polymerase III (Werner, M., personal communication and Mann *et al.,* 1992). Since the two-hybrid interactions with Ymr44p and Rpc53p are the only information on Ybr223p available to date, the

significance of these interactions is difficult to evaluate. The protein will not be discussed further for that reason.

YNL078w encodes a hypothetical 46 kDa protein, which shares no homologies to other proteins in the databases. The protein is basic (pI 10.16) and unusually serine rich (15% of all amino acids). The common region of the prey inserts consists of the carboxyterminal 77 amino acids (see figure IV.13).

The *YHR134w* ORF codes for a 30 kDa hypothetical protein, with no homologies to other proteins in the databases. The c-terminal half of the protein seems to be sufficient to interact with the Ymr44p bait (figure IV.13).

In summary, Ymr44p was found to interact with the second step splicing factor Prp18p in a two-hybrid screen and when it was used in turn as bait in a two-hybrid screen it interacted with the polyadenylation protein Fir1p. Thus, the protein could function at a putative splicing/polyadenylation interface. Additional information yielded in two-hybrid screens performed within and outside the TAPIR network do indeed support this idea. These interactions will be discussed in detail in the next section.
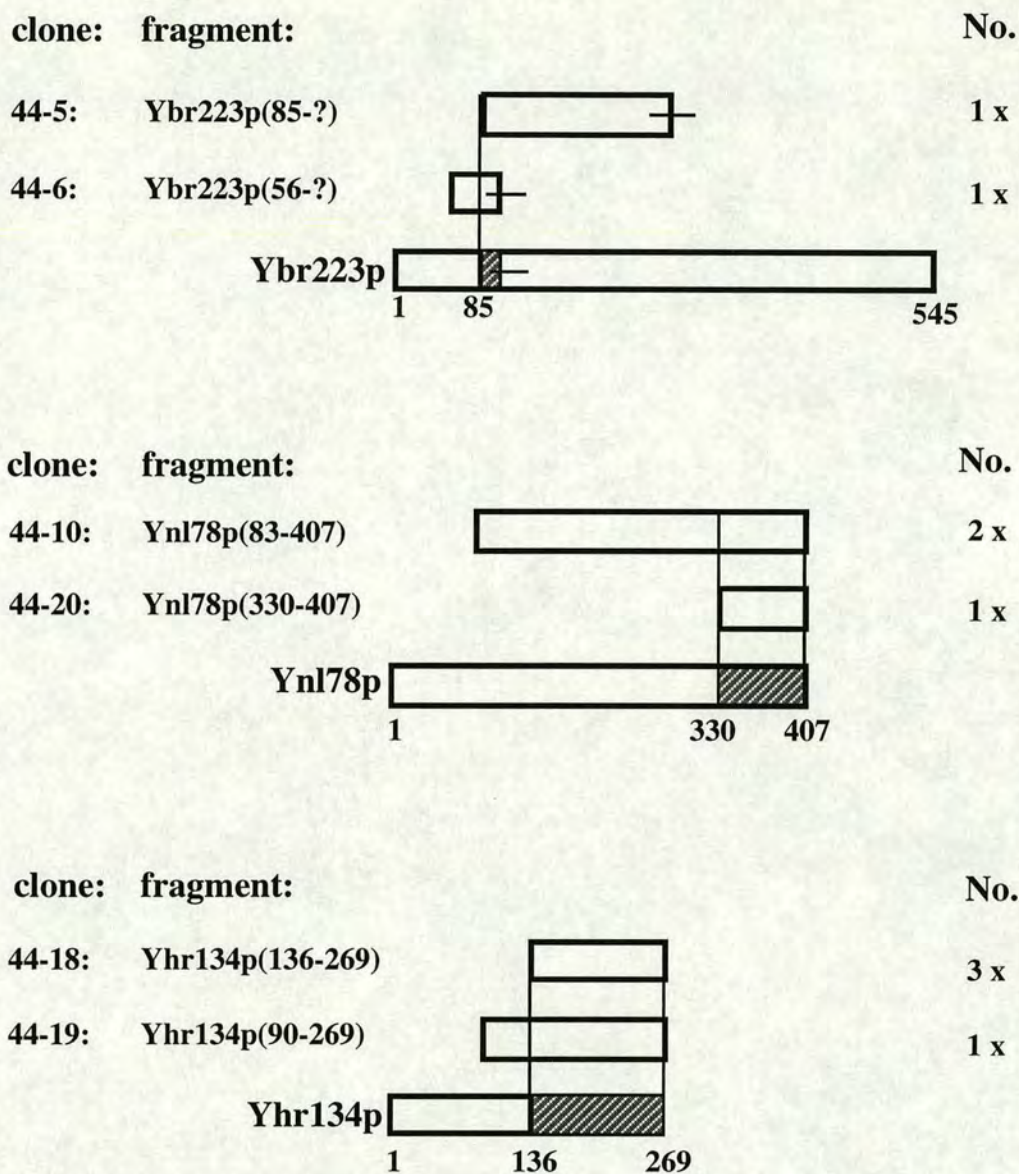
Figure IV.13:
A1 category preys identified in the Ymr44p two-hybrid screen.
Besides Fir1p (figure IV.12), three additional A1 preys were identified with the LexA::Ymr44p bait: Ybr223p, Ynl78p and Yhr134p. The prey fragments which were found are depicted here. For the Ybr223p prey fragments, the C-terminal end point was only estimated from the size of the inserts in the prey plasmids, determined via *Bam*HI digest and subsequent agarose electrophoresis). This is indicated by a thin "error line" at the end of the boxes depicting the fragments. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets define the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen.

IV.3.4 Discussion

In this work, the Ymr44 protein was shown to interact in a two-hybrid screen with the second step splicing factor and U5 snRNP-associated protein Prp18p (Horowitz and Abelson, 1993). Additional two-hybrid screens performed within the TAPIR network using other splicing factors as baits also identified the protein as interactor: Ymr44p prey was identified in two-hybrid screens with the U1 snRNP-associated proteins Snp1p (Neubauer *et al.*, 1997) and Yhc1p, the yeast U1C homologue (Tang *et al.*, 1997) and in addition with the U2 snRNP-associated factor Rse1p (Caspary *et al.*, 1999; Fromont-Racine *et al.*, 1997 and Pierre Legrain, unpublished results). It is questionable, whether all (or any one) of these interactions are (is) specific and functionally relevant. It may be not too surprising to find a protein interacting with U5 snRNP proteins (Prp18p) as well as U1 proteins (Snp1p and Yhc1p), because the proximity of the U1 and U5 snRNPs and snRNAs within the spliceosome has been noted before (Gottschalk *et al.*, 1998; Ast and Weiner, 1997). However, since Prp18p acts at the stage of 3'-splice site recognition and cleavage, whereas Snp1p and Yhc1p both are involved much earlier, during commitment complex formation, it is unlikely that Ymr44p interacts with all three proteins. The same argument holds for a putative Rse1p-Ymr44p interaction. Rse1p has been proposed to be the homologue of the human splicing factor SF3b[130] (Caspary *et al.*, 1999), since the proteins share extensive similarities in sequence and function. No interactions of the SF3b[130] protein with human second step splicing factors have been reported so far. Thus, it is unlikely that Ymr44p interacts with both Prp18p and Rse1p during the splicing process.

Another point, which has to be noticed, is that the above mentioned two-hybrid interactions are not very strong with respect to the number of fusions found in the screens. In the Snp1p-, Yhc1p- and Rse1p two-hybrid screens, Ymr44p was an A4 interactor. In case of the Prp18p interaction, *two* screens had to be done to find two independent fusions of the Ymr44p. The rarity of the interactions might reflect the fact that they (or some of them) are aspecific random interactions. This assumption is supported by the results of the Ymr44p two-hybrid screen. None of the interactions discussed above could be confirmed in the reciprocal orientation, nor were any other

splicing factors found with the Ymr44p bait. Instead, the results of the Ymr44p screen suggest a close link of the protein to the polyadenylation machinery. Two findings underline this hypothesis:

First, the strongest interactor found, Fir1p, is a protein required for efficient polyadenylation. Fir1p was also found in two-hybrid screens with other polyadenylation factors, namely as A1 interactor with Ref2p (Russnak *et al.*, 1996), as A1 interactor with the PolyA polymerase (Pap1p, del Olmo *et al.*, 1997) and as A1 candidate with Hrp1p. Hrp1p is the CF IB component of the cleavage factor CF I (Kessler *et al.*, 1997), which has been shown to be required for accurate cleavage site choice, but not for cleavage per se (Minvielle-Sebastia *et al.*, 1998).

Second, two of the three additional A1 prey proteins found with the Ymr44p bait, namely Yhr134p and Ynl78p, were identified in two-hybrid screens with polyadenylation factors done within the TAPIR network (M. Minet and co-workers, unpublished results), suggesting a function of these proteins during 3'-end formation of the message. Both proteins represent A4 candidates in a Hrp1p screen and furthermore, Ynl78p was found as A1 interactor with Rna14p, a component of the cleavage factor CF 1A (Minvielle-Sebastia *et al.*, 1994). An overview of the numerous two-hybrid interactions discussed in this section is given in figure IV.14.

In conclusion from the two-hybrid data available, it is most likely that Ymr44p is involved in polyadenylation of the messenger RNA. The protein is not essential and it could therefore be imagined, that it may have a modulating function like e.g. Fir1p and Ref2p. The two-hybrid data which link Ymr44p to the splicing machinery are somewhat diffuse and it is difficult to suppose at what step of the splicing process the protein could act. Moreover, the weakness of the two-hybrid interactions to the splicing factors might suggest that these interactions are random and aspecific. Further experiments need to be done to establish whether there indeed is a role for this protein in splicing and whether Ymr44p connects the two RNA processing events.

Figure IV.14:
Two-hybrid network illustrating the interactions of Ymr44p with the splicing- and polyadenylation machinery.
Known splicing factors are depicted in orange boxes, polyadenylation proteins in blue boxes and proteins of unknown function in yellow boxes. Bold arrows represent A1-, thin ones A4-interactions.* two independent fusions found in *two* screens.

Symbols indicate who performed the screen: $ - Fromont-Racine *et al.*, 1997; Pierre Legrain, unpublished results; % - Micheline Minet, unpublished results; & - Russnak *et al.*, 1996; @ - del Olmo *et al.*, 1997; r - Rob van Nues, this laboratory; £ - this work.
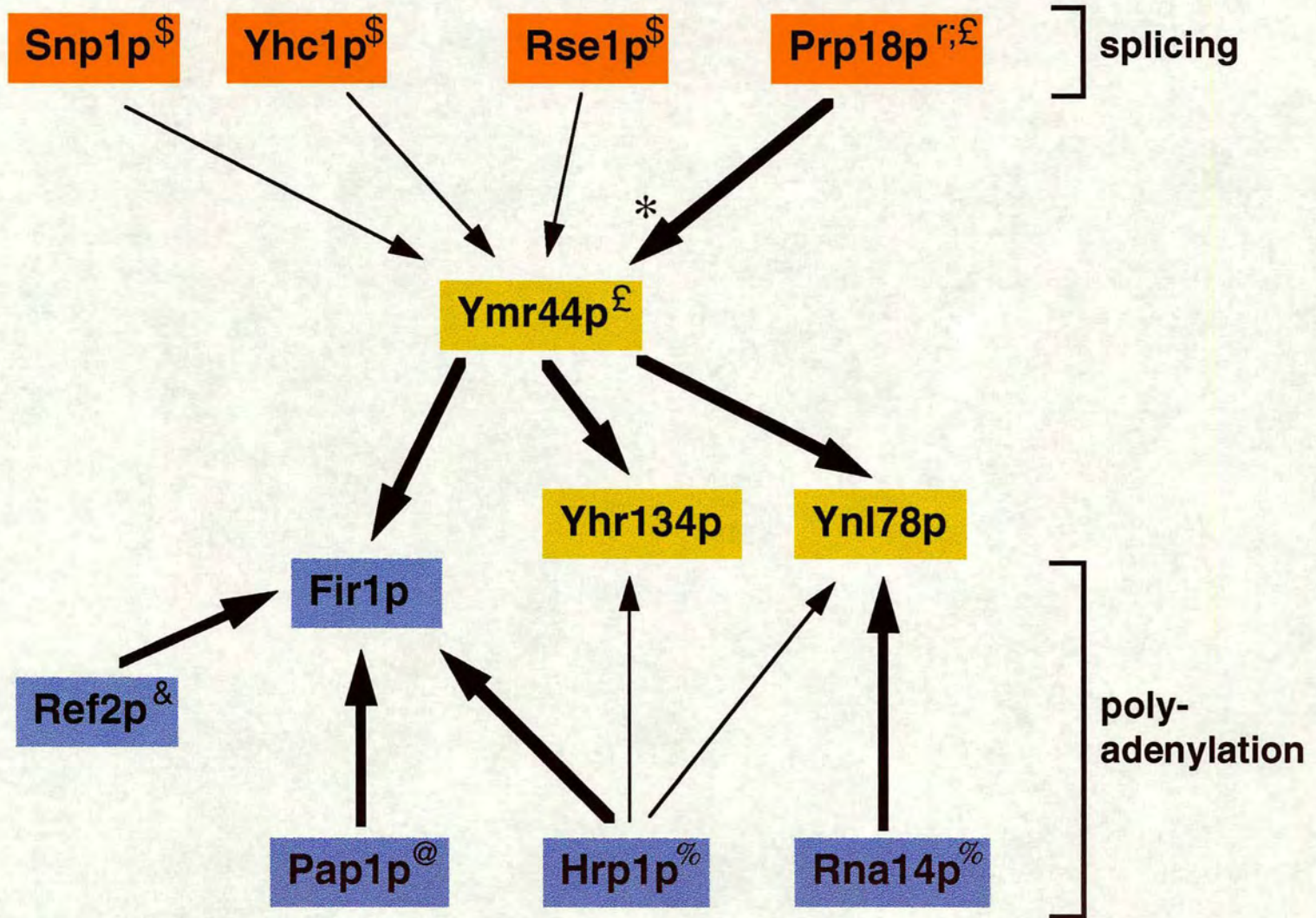
Chapter V

The Prp22p two-hybrid screen

V.1 Introduction

The interesting finding that Prp22p fragments interacted with the second step splicing factor Slu7p in a two-hybrid screen suggested that the two proteins might interact with each other in the course of 3' splice site selection and cleavage. In order to investigate this hypothesis and possibly to reveal additional interactions, Prp22p was used as a bait in a "second round" two-hybrid screen. The performance and the results of the Prp22p two-hybrid screen will be described in this chapter.

V.2 Construction of the *PRP22* bait vector

The full length *PRP22* ORF was generously provided on plasmid pBSKS22 by Beate Schwer (table II.1.8). This plasmid had been generated by cloning a *Sac*I, *Kpn*I fragment of the *PRP22* locus into the polylinker of plasmid pBluescript II KS+ (Stratagene). An *Nde*I restriction site had been engineered into the start site of the *PPR22* ORF, in such way that the terminal three nucleotides of the *Nde*I recognition site (CAT**ATG**) represent the translational start codon of the *PRP22*-ORF. This allowed an easy transfer of the ORF into other vector systems.

The plasmid pBSKS22 was cut with *Kpn*I, blunted by T4 DNA polymerase I treatment and then cut with *Nde*I. The fragment containing the *PRP22* ORF was separated from the vector via agarose gel-electrophoresis and isolated from the gel using the QIAquick Gel Extraction Kit (Qiagen). Subsequently it was ligated into the pAS2ΔΔ vector, which had been cut with *Nde*I and *Sma*I to generate the *PRP22* bait vector pMA22AS. The identity of the vector and the correct frame of the *GBD::PRP22* fusion was checked by DNA-sequencing using oligonucleotides N3027 (pAS2ΔΔ forward) and P5148 (pAS2ΔΔ reverse). The pMA22AS vector was transformed into the yeast strain CG1945 and crude cell extracts were prepared from the cells. The production of the Gal4p DNA-binding domain-Prp22p fusion protein (GBD::Prp22p) was then checked by western blotting using antibodies against the GBDp. A protein of the expected size of 144 kDa could be detected (data not shown).

In a directed two-hybrid mating assay in which the Prp22p bait was checked for interactions with 19 random preys that were available in this laboratory, no interactions could be detected, showing that the bait fusion did not strongly autoactivate (data not shown). Thus the Prp22 bait was ready to be used in a two-hybrid screen.

V.3 The Prp22p two-hybrid screen

A full scale two-hybrid screen using the GBD::Prp22p fusion protein as bait was performed using 1 vial of the FRYL. 47 million diploid yeast clones were screened for expression of the *HIS3* reporter gene by plating them onto YMM -LWH plates. The mating efficiency was calculated as 20 %. After three days, the plates were overlaid with an X-gal agar mix, and 26 colonies turned blue by 10 hours of incubation. No additional blue colonies could be spotted at later time points. When the 26 colonies were streaked onto fresh YMM -LWH agar, 1 colony did not regrow. After restreaking the cells another two times, a β-gal filterlift assay was performed. Five clones did not develop a blue colour in this assay and were discarded. The identity of the library inserts in the prey plasmids of the remaining twenty clones was determined: a yeast colony PCR was performed to amplify a region of the prey vector containing the insert. Oligonucleotides T3785 and T3786 were used for that purpose. The amplified inserts were subsequently gel purified and sequenced, using the same oligonucleotides (T3786 for forward-, T3785 for reverse sequencing). The positive candidates of the Prp22p two-hybrid screen are listed in table V.1, which also includes the relative strength of the observed interaction in the filterlift assay in column "interaction".

Table V.1:
Results of the Prp22p two-hybrid screen. A Gal4 DNA-binding domain-Prp22 fusion protein (GBD::Prp22p) was used as bait to screen the FRYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. No. - frequency with which a fragment was identified in the screen; Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - insert length (nucleotides) determined by DNA sequencing; Interaction - relative strength of response in β-gal filterlift assay: +++ very strong response; ++ strong response, + moderate response; preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size (bp)	Insert size (bp)	Interaction	Category	Protein info
22-1	5	*FUN20*	*YAL032c*	I	c	-23	1140	1073	+++	A1	**Prp45p, splicing factor (this work)**
22-19	1	*FUN20*	*YAL032c*	I	c	632	1140	1243	+++	A1	**Prp45p, splicing factor (this work)**
22-10	2	*FUN20*	*YAL032c*	I	c	25	1140	849	+++	A1	**Prp45p, splicing factor (this work)**
22-27	1	*FUN20*	*YAL032c*	I	c	782	1140	298	+	A1	**Prp45p, splicing factor (this work)**
22-13	1	*FUN20*	*YAL032c*	I	c	580	1140	831	++	A1	**Prp45p, splicing factor (this work)**
22-23	1	PMI40	YER003c	V	c	1024	1383	500	++	A4	Mannose-6-phosphate isomerase
22-14	2	DBP7	YKR024c	XI	c	919	2229	882	++	A4	RNA-helicase, 60S ribosomal large subunit
22-12	3	*SYF3*	*YLR117c*	XII	c	1765	2064	377	++/+	A1	**Splicing factor**
22-22	1	*SYF3*	*YLR117c*	XII	c	1862	2064	528	++	A1	**Splicing factor**
22-29	1	*SYF3*	*YLR117c*	XII	c	1880	2064	1381	+	A1	**Splicing factor**
22-25	1		YNL274c	XIV	c	97	1053	1113	+	A3	
22-20	1	anti	(YDR409w)	IV	c	-20	2715	1188	++	B	

Two candidates were found to interact with the Prp22p bait that fall into the A1 category: The 42 kDa Fun20 protein (referred to as Prp45p, see below), encoded by ORF *YAL032c*, was identified 10 times in 5 different fusions. The fragments found with the Prp22p bait are depicted in figure V.1 A. The putative interacting region, the overlap between the prey fragments spans the amino acid residues 262-291 of the 379 amino acid long protein. It might be worth noticing that this small region of 30 amino acids contains two putative phosphorylation sites, which are also indicated in figure V.1 A.

The protein was known to be essential for cell viability (Diehl and Pringle, 1991), but no additional data about a putative function were available. Fun20p was extensively studied in this work and it has been demonstrated that it is a novel pre-mRNA splicing factor. It was therefore renamed and henceforth will be referred to as Prp45p in this work. It will be discussed in detail in chapter VI.

The second A1 prey candidate found was the 82 kDa Syf3 protein, encoded by ORF *YLR117c*. It was found 5 times in 3 different fusions, which all share the c-terminal 59 amino acids of the protein (figure V.1 B).

Originally, the *SYF3* gene was isolated in a screen searching for mutants synthetic lethal with a deletion of the gene for the second step splicing factor and cell cycle protein Prp17p, also named Cdc40p (Ben-Yehuda and Kupiec, unpublished results).

The Syf3 protein shows homology to the putative cell cycle regulatory protein crooked neck (Crn) of *Drosophila melanogaster* and was therefore alternatively named Clf1p (for crooked necked like factor). It contains so called tetratricopeptide (TPR) motifs that might be involved in protein-protein interactions (for a review see Lamb *et al.*, 1995; see also chapter VI.10.3 for a more detailed discussion of the TPR repeats of Syf3p). Recently it was demonstrated that it is a splicing factor required for the first transesterification to occur and it was proposed to facilitate tri-snRNP addition during spliceosome assembly (Chung *et al.*, 1999). The Syf3 protein was also found as prey of the A1 category in a Prp45p two-hybrid screen (see chapter VI).

A)

clone: fragment: No.

1:         Prp45p(1-350)   5 x

19:        Prp45p(212-379)   1 x

10:        Prp45p(9-290)   2 x

27:        Prp45p(262-359)   1 x

13:        Prp45p(194-379)   1 x

           Prp45p

           1   262 291   379
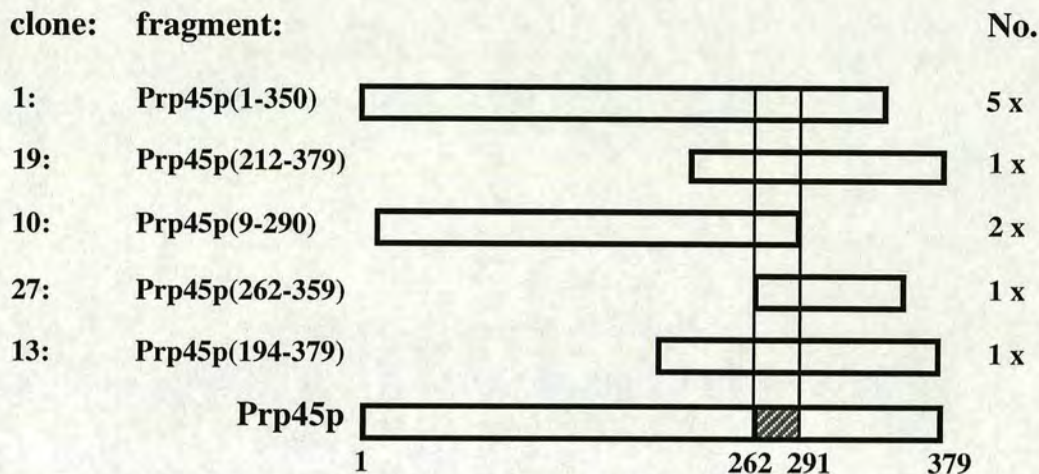
common domain: 262 KP KKQT STVA RLKELAY SQGRDVSEKIILG 291

B)

clone: fragment: No.

12:        Syf3p(592-687)   3 x

22:        Syf3p(622-687)   1 x

29:        Syf3p(628-687)   1 x

           Syf3p

           1   628   687
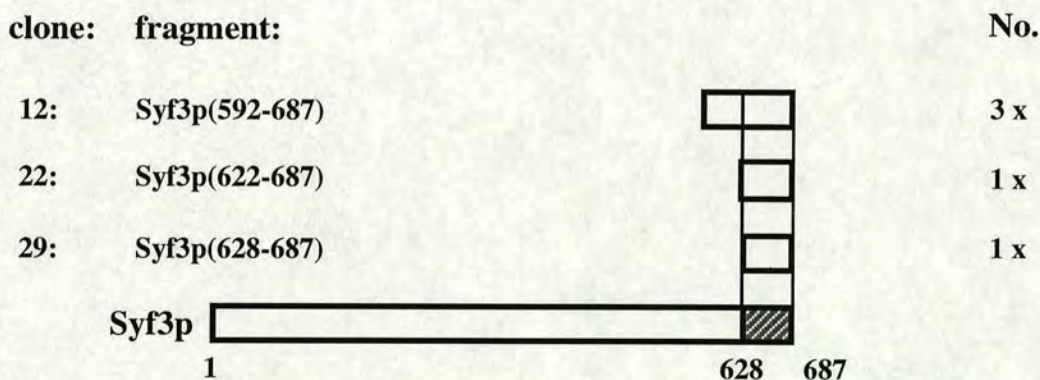
Figure V.1:
Prey fragments of Prp45p and Syf3p found in the Prp22p two-hybrid screen. The common domain of the fragments is presented as a shaded box. The numbers in brackets indicate the amino acid boundaries of the particular fragment. No. - frequency with which a particular clone was found in the screen. **A)** Prp45p prey fragments. The primary structure for the Prp45p common domain is given and amino acid residues representing two putative phosphorylation motifs are boxed. Threonine (T) and tyrosine (Y) residues that are potential targets for phosphorylation are highlighted in bold type. **B)** Syf3p prey fragments.

The protein encoded by ORF *YNL274c* was found as an A3 interactor with Prp22p. A strong sequence homology of the protein to dehydrogenases in a range of organisms from archaebacteria to higher eucaryotes suggests a function unrelated to

pre-mRNA splicing. For that reason the protein was not considered for further investigation.

The two A4 prey candidates found are both well characterised: Pmi40p is the mannose-6-phosphate isomerase (Smith *et al.*, 1992) and Dbp7p is a putative RNA-helicase involved in ribosomal biogenesis (Daugeron and Linder, 1998).

V.4 Discussion

In the two-hybrid screen with the Prp22p bait, two splicing factors have been found as A1 candidates: the Fun20 protein, which is subsequently referred to as Prp45p, and Syf3p. Prp45p was extensively studied in this work and its characterisation and the significance of its interaction with Prp22p will be subject of the next chapter(s).

Syf3p, the other A1 prey candidate found with Prp22p, was also found as an A1 prey in a two-hybrid screen with Prp45p as bait. Therefore, the Syf3 protein and the implications of the triangular two-hybrid interaction between the three proteins, will be discussed after the characterisation of Prp45p in chapter VI.

The amino acid sequences of the proteins Ynl274 (a putative dehydrogenase) and Pmi40 (mannose-6-phosphate isomerase), which were found as A3 and A4 candidates respectively, do not give any clue as to why these proteins have been identified as likely false positives with Prp22p. For the Dbp7p, a DEAD-box protein involved in ribosomal biogenesis, however, it is possible that the interaction with Prp22p in the two-hybrid assay is due to the sequence homology of the protein with the DEAH-box protein and second step splicing factor Prp16p. As discussed in chapter III.3, in a two-hybrid screen using the Prp16 protein as bait, Prp22p was found as A1 prey candidate. In fact Prp22p was the only interacting protein found in this screen (Alan Colley, personal communication). The 294 amino acid long prey-fragment of Dbp7p, that was found to interact with Prp22p, comprises a very conserved region within the family of DEAD/H-box proteins, containing the conserved motifs 2-6 (for a recent review on DEAD-box proteins see Linder *et al.*, 1999). The homology to Prp16p in this region is

rather small (only about 15% amino acid identity), but there might be common features within this region of Prp16p and Dbp7p, that led to the isolation of Dbp7p with the Prp22p bait. However, the question must be asked why Prp16p itself was not found with the Prp22p bait, but instead a distant member of the DEAD-box family, such as Dbp7p. It has been found for many two-hybrid interactions, that even in cases where the biological significance of the interaction is obvious, a reciprocal interaction could not be observed. Problems with the folding of particular fusion proteins or with instability of the fusion proteins could contribute to this behaviour.

As for the Prp22p-Prp16p interaction, the interaction of Slu7p with Prp22p, observed in the two-hybrid screen Slu7-III (chapter III.2.2), was not found in the reciprocal orientation. Therefore, the Prp22p screen did not put further weight to the hypothesis that Prp22p contacts the second step splicing factors Slu7p or Prp16p in the course of 3' splice site selection and cleavage.

Chapter VI

Prp45p is a novel pre-mRNA splicing factor

VI.1 Introduction

The Prp45 protein was found as the statistically most significant prey in a two-hybrid screen using the second step- and spliceosome disassembly factor Prp22p as bait. Prp45p has a calculated molecular weight of 42456 Da and consists of 379 amino acid residues. The protein is hydrophilic; 30% of its amino acid residues are charged. Overall, Prp45p is of basic character with a predicted isoelectric point (pI) of 9.29. The charged residues are spread over the protein, with no large stretches of consecutive acidic or basic residues (see also figure VI.1 for a summary of primary structure features of the protein).

Prp45p was already known to be essential for cell viability (Diehl and Pringle, 1991), but no additional functional data were available. In this chapter the first evidence will be presented that Prp45p is a pre-mRNA splicing factor.

VI.2 Prp45p has sequence homology to proteins in other eukaryotes

An NCBI blast protein database search revealed that the Prp45 protein shares significant sequence homology to proteins in a range of eucaryotes from plasmodium to man. These proteins show a particularly strong conservation of primary structure in an n-terminal/central region spanning approximately 170 aminoacid residues. No function has yet been assigned to this conserved region, which was previously named SNW-domain, paying tribute to a 100% conserved motif consisting of the amino acids SNWKN, present in all members of the group found to date. Figure VI.2 shows a sequence alignment of the conserved proteins.

Figure VI.1:
Primary structure analysis of Prp45p.
A) Schematic illustration of the positions of acidic (A) and basic (B) amino acid residues within the 379 amino acid long protein. **B)** Hydrophobicity plot after Kyte and Doolittle (1982) (windows of 11 amino acids were used for calculation). Positive values represent hydrophobicity, negative values hydrophilic stretches. **C)** Amino acid composition: n - number of times a particular amino acid is present in the protein; MW - molecular weight. (the above analyses were done using the DNA strider 1.2 (GATC) program)

A)
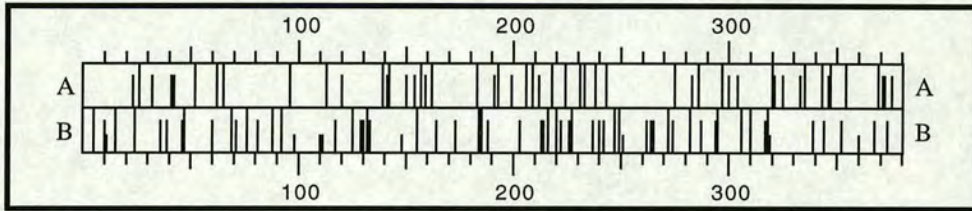
B)
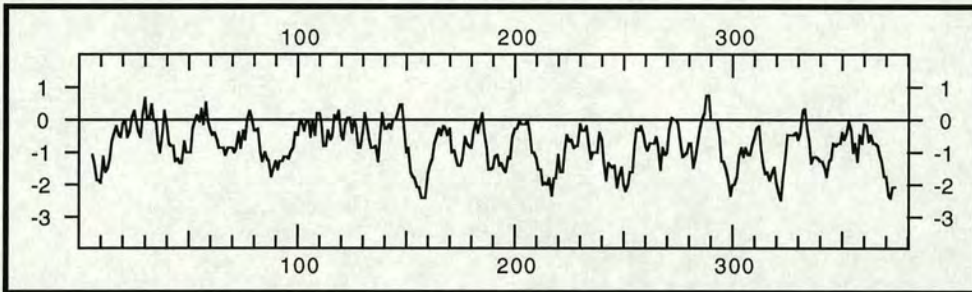
C)

```
        379 Amino Acids   MW :    42456 Dalton

                            n    n(%)        MW   MW(%)

A ala  alanine             35    9.2        2486    5.9
C cys  cysteine             1    0.3         103    0.2
D asp  aspartic acid       22    5.8        2530    6.0
E glu  glutamic acid       29    7.7        3742    8.8
F phe  phenylalanine        9    2.4        1323    3.1
G gly  glycine             16    4.2         912    2.1
H his  histidine            8    2.1        1096    2.6
I ile  isoleucine          15    4.0        1696    4.0
K lys  lysine              34    9.0        4355   10.3
L leu  leucine             29    7.7        3279    7.7
M met  methionine           7    1.8         917    2.2
N asn  asparagine          27    7.1        3079    7.3
P pro  proline             20    5.3        1941    4.6
Q gln  glutamine           19    5.0        2433    5.7
R arg  arginine            24    6.3        3746    8.8
S ser  serine              31    8.2        2697    6.4
T thr  threonine           17    4.5        1717    4.0
V val  valine              24    6.3        2377    5.6
W trp  tryptophan           2    0.5         372    0.9
X ukw  unknown              -    -
Y tyr  tyrosine            10    2.6        1630    3.8
Z ---  STOP                 -    -
```

Figure VI.2:
Sequence alignment of Prp45p with putative homologues in other eukaryotes. Proposed homologues were identified by BLAST searches of the NCBI nr peptide sequence database (section II.4). Sequences were aligned using the PILEUP program of the GCG9 suite of sequence analysis programs, with identities and similarities highlighted using BOXSHADE 3.21 (section II.4). Black boxes highlight identical residues, grey boxes indicate the conservation of the nature of the amino acid at that site. Accepted conservative groupings were M=I=L=V, K=R=H, F=Y=W, S=T, E=D, A=G and Q=N. A particular amino acid position is boxed if at least 6 of 8 residues at that position are identical or belong to the same conserved group. A dashed box indicates a conserved proline (P)-rich motif, the grey box the "SNW domain" (see text), which includes the 100% conserved amino acid residues SNWKN (highlighted by a black box within the grey box). The location of regions with high similarity to structural motifs of SH2-domains (helix A, β-strand B, β-strand C and β-strand D) is indicated by open boxes. Gene names or accession numbers are given in brackets; *H. s - Homo sapiens*, *D. m - Drosophila melanogaster*, *C. e - Caenorhabditis elegans*, *A. t - Arabidopsis thaliana*, *S. p - Schizosaccharomyces pombe*, *D. d - Dictiostelium discoideum*, *P. f - Plasmodium falciparum*, *S. c - Saccharomyces cerevisiae*. (note that the order of the *H. s* and *D.m* proteins has changed from the first to the second sheet of the alignment).

```
H.s (SNW1)        1  ~~~~~~MALTSFLPAPTQLSQDQLEA...EEKA.RSQRSRQTSLVSSR....R.EPPPYGYRKGWIFRLLEDFGDGGAFPEIHVAQYPLDM...GR
D.m (Bx42)        1  ~~~~~MSLSSLLPTPTNAIWDREDE...RRLV.ARGAPKIGALVSAK....I.AAPPYGQRKDWVPHTDADFGDGGAFPEIHVAQYPLGLGAPGN
C.e (Q22836)      1  ~~~MSMKLRDILPAPVAADEAASQI...RRDPWFGGRDNEPSAALVS....K.EPPPYGKRTSFRPRGPEDFGDGGAFPEIHVAQFPLGLGLGDM
A.t (gi3540201)   1  ~~~~MKSLNDLPAPKSTTTTYYDH...SNDAWFKNRVTESETVKSSSIKFK.VVPAYLNRQGLRPKNPEDFGDGGAFPEIHLPQYPLLMGKNKS
S.p (Q09882)      1  MALLSEELSSILPNP....DFDDE...EEDYV....ERETSHADERQIGVKFHIPPYGQRKGWFPSSPEDFGDGGAFPEIHVAQYPLDMGRKRS
D.d (SNWA)        1  ~~~MTSLSSLLPKPKNVYSNEEEDPLFQPKPKPQQQQKQQQQQQQELNDKPKKVIPTYGNRKGYLPKNIEDFGDGGAFPEIHIVQYPLDMGRKGK
P.f (gi3845299)   1  ~~~~MTDFLRNIPKPKKKAYDDENELHD......FKESNNSIKKKEEIKKKNQCYEYLKRRHLRITCNEDFQGGAVPEIHMNQYPNNIGLKSD
S.c (Prp45p)      1  ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

H.s (SNW1)       79  KKKM.SNA...........LAIQVDSEGKIKYDAIARQGQSKDKVIYSKYTDLVPKEV.MNADDPDLQRPDEE....AIKEITEKTRVALEKSV
D.m (Bx42)       82  VGKK.SDA...........LAVRLDDKGKVKYDAIARQGHGKDKIVYSSISQLLPAEV.LAEDADELQRPDEE....TVMETTEETRLALEKLT
C.e (Q22836)     85  RGKP.ENT...........LALQYGTDGKLQHDAIARIGHVKDKVVYSKLNDM.KAKT.WNEDDDDIQKPDDD....AVIDATEKTRMALEKIV
A.t (gi3540201)  87  NKPG.AKT...........LPVTVDAQGNVVFDAIVRQNENSRKIVYSQHKDIIPKFL.KNEGDLGTVVDEEEELQKEIQETAEETKAAIEKIV
S.p (Q09882)     84  AKSA.GNT...........LALQVTSSGAVDYNAIARQGHEHGELVQASFRDLIPLRARLGVGEISLEKPSDEQKQ....EVANKTKLALQKIL
D.d (SNWA)       92  SKSSNSNTSNMNGGGTTTSIVPVSVDSTGRVKHEAIL....GEKGSLHSQYKDLIPKQ....HTEHELQRPDDDELQ....ETLDRTKNALEKIV
P.f (gi3845299)  85  NKNNIVLKYIDENNNVKYDNL...INQQIHIYNNEIDKIEPNERINKLRKKKILSDTKDREEKYNEPTYKPNNDEEN....EIIENTKRNIENIL
S.c (Prp45p)      1  ~~~~~~~~~~~~~~~~MFSNRLPPPKHSQGRV..STALSSDRVEPAILTDQIAKNVKLDDFIPKRQSNFELSVPLPTKAEIQECTARTKSYIQRLV

H.s (SNW1)      156  SQKVAAAMPVRA.ADKLAPAQYIRYTPSQQGVAFNSGA.KQRVIRMVEMQKDPMEPPRF...KINKKIPRGPPSPPAPVMH.SPSRKMTVKEQQE
D.m (Bx42)      159  NQKITSALPVRH.AQKAGPAQYIRYTPSQQGDTFNSGA.KQRVIRMVEAQLDPMEPPKF...RINKKIPRGPPSPPAPVLH.SPSRKVTVKEQKE
C.e (Q22836)    161  NSKVASALPVRH.ADKLAPAQYIRYTPSQQNGA..AGS.QQRIIRMVEEQKDPMEPPKF...KINQKIPRAPPSPPAPVMH.SPPRKMTAKDQND
A.t (gi3540201) 168  NVRLSAAQPSNI.ARQSGDSQVIVKPSQQSSAFNSGA.KERIIRMVEMPVDPLDPPKF...K.HKRVPRASGSPPVPVMH.SPPRPVTVKDQQD
S.p (Q09882)    162  SKQIAQSQPKSAVVQQRDDPVYIRYTPSNQMG..QALS.KQRIIKMVTAEQDPMEPPKF...R.HKKVPRGPPSPPVLH.SPPRKVSAQEQQD
D.d (SNWA)      175  NGKIKSSKSTNYVEVEKKSATYIKYTPSNQLGSNNGSALNSKIVRMVDVAQDPLEPPKY...KIKKKIME.HGSPPAPVMH.SPTRKLSVQDQQD
P.f (gi3845299) 173  NEKLNKSNIVNKKEEK.....YYRYIPQNKLN....NNLEERIIKIVEKGTDPLDVSKF...K.HKKLPNIKNSPDYPILR.SPTRKLNKEEEND
S.c (Prp45p)     79  NAKLANS.......NNRASSRYV..TETHQAPA.NLLLNNSHHIEVVSKQMDPL.LPRFVGKKARKVVAPTENDEVVPVLHMDGSNDRGEADPNE

H.s (SNW1)      245  WKIPPCISNWKNAKGYTIPLDKRLAADGRGLQTVH..INENFAKLAEALYIADRKAREAVEMRAQVERKMAQKEKEKHEEKLREMAQKARERRAG
D.m (Bx42)      248  WKIPPCISNWKNAKGYTIPLDKRLAADGRGLQQVH..INEKFAKMAEALYIADRKAREAVEARSQLEKKLAQKEKEKKEDMLRMMAQRAREERAG
C.e (Q22836)    248  WKIPPCISNWKNPKGFTVGLDKRLAADGRGLQQTH..INENFAKLADALYIADRKAREEVETRAQLERRVAQNKKSEQEAKMAEAAAKARQERSA
A.t (gi3540201) 256  WKIPPCISNWKNPKGYTIPLDKRLAADGRGLQDVQ..INDNFAKLSEALYVAEQKAREAVSMRSKVQKEMVMKDKERKEQELRALAQKARSERTG
S.p (Q09882)    249  WQIPPSISNWKNPKGYTIPLDKRLAADGRGLNDVE..INDGFAKFSEALYTVERQAREEVRYRAIMRQKMAEKEKQEKEQRLFMLAQKAREDRMG
D.d (SNWA)      265  WTIPPCVSNWKNPKGFAISIDKRLVSNGGGLQDVE..INDKFAHFTQALYIAESNAREEVSARAELERKLAQKEKERKQDMLRKLAEDVRNERSG
P.f (gi3845299) 254  WKIPPCVSNWKNNKGYNIPLKRIQSDNKKLNNVV..VNENFAHLSEYLYVAEKKAREEIQIRNSVMKQKKLKEEKEENVLKNLAIQARKEK.G
S.c (Prp45p)    163  WKIPAAVSNWKNPNGYTVALERRV...GKALDNENNTINDGFMKLSEALENADKKARQEIRSKMELKRLAMEQEMLAKESKLKELSQRARYHNGT
```

continued...

```
D.m (Bx42)        341 ------------------------------------------------------------------------------------
H.s (SNW1)        338 ------------------------------------------------------------------------------------
C.e (Q22836)      341 ------------------------------------------------------------------------------------
A.t (gi3540201)   349 ------------------------------------------------------------------------------------
S.p (Q09882)      342 ------------------------------------------------------------------------------------
D.d (SNWA)        358 IIQQRYTRKDNSDSDNDNDNDSSSDEDKNKRTPPMNRRSRSRSTERIPSRNDNDDDDDRYRIKDNRDNRGRDNIDSRDNRDSRDSRDSRDSR
P.f (gi3845299)   346 ------------------------------------------------------------------------------------
S.c (Prp45p)      255 ------------------------------------------------------------------------------------


D.m (Bx42)        341 ---------------------------------------LRNPEAAEPSGSGATGSEVRERNDLRAERQRERQRDRNLQR..AAPEKR
H.s (SNW1)        338 --------------------------------IKTHVEKEDGEARERDEIRHDRRKERQHDRNLSR..AAPDKR
C.e (Q22836)      341 --------------------------MRRKDDEDDEQVKVREEIRRDRLDDIRKERNIAR..SRPDKA
A.t (gi3540201)   349 -----------AAMSMPVSSDRGRSESVDPRGDYDNYDQDRGREREREEPQETREEREKRIQREKIREEERRRERERERRLDAKDAAMGKK
S.p (Q09882)      342 ------------RNAASSGPSHAKPRSTSVSSEERSRSRAGSFSHHSESENEDEDSEAFRRRQELRRERRRQAEKDLRLSRMGA..EKR
D.d (SNWA)        453 DSRDSRDNRDSRDSRDNRDNRDNRRRDDSNDRDRYSKRRSDSDSDSDSDSSDSEDERVRRERKEKLERDKIRMEKKRELEREYRLE....ASGKK
P.f (gi3845299)   346 ---------------------------------------LAHSSLINDRKREIEREYKI......EKN
S.c (Prp45p)      255 -----------------------------------PQTGAIVKPKKQTSTVAR

                                      helix A                        β—strand B      β—strand C                           β—strand D
D.m (Bx42)        388 SKL.QKERERDISEQIALGLP.AKSAGNGETLFDQRLFNTTKGMDSGY..GDDEAYNVYDKPWRDSNTLGAHIYR.PSKQADSDNYGG...DLDA
H.s (SNW1)        378 SKL.QRNENRDISEVIALGVP.NPRTSN.EVQYDQRLFNQSKGMDSGFAGGEDEIYNVYDQAWRGGKDMAQSIYR.PSKNLDKDMYGD...DLEA
C.e (Q22836)      381 DKL.RKERERDISEKIVLGLPDTNQKRTGEPQFDQRLFDKTQGLDSGAM..DDDTYNPYDAAWRGGDSVQQHVYR.PSKNLDNDVYGG...DLDK
A.t (gi3540201)   428 SKI.TRDRDRDISEKVALGMASTGGKGGGEVMYDQRLFNQDKGMDSGFAA..DDQYNLYDKGLFTAQPTLSTLYK.PKKDNDEEMYGNADEQLDK
S.p (Q09882)      416 AKLAEKDRPRDVAERVALGL..SKPSMSSDTMIDSRLFNQASGLGSGFQ..DEDSYNVYDKPWRAAPS..STLYR.PGATLSRQV..DASAELER
D.d (SNWA)        544 SKF.NRDQDRDISEKIALGQASIKR..TEDSIYDQRLFNQSESLTSGF.GNDDSYNVYSKPLFGG.AVSNSIYR.PKSNQEDNTSIQDVLSNSR
P.f (gi3845299)   369 LKKMKNYENRYVEEQIALNKVNVSKNNN...IHDITLFNINE.QNNVTTTQDDDTYQIYDTALFNNKN.NANIYKFSSERLRKNVQKIETRDTMQ
S.c (Prp45p)      273 LKELAYSQGRDVSEKIILG..AAKRSEQPDLQYDSRFF..TRGANASAKRHED...QVYDNPLFVQQDI.ESIYKTNYEKLDEAVNVKSEGASGS


D.m (Bx42)        475 IV.N..TKRFV...PDKQFSGASKEAAAGQRSGPVEFEKEE....DPFGLDQFLNMAKK..APKRAEEKNNERS.SHSDRKRSKRD~~~~~~
H.s (SNW1)        466 RI.K..TNRFV...PDKEFSGSDRRQRG..REGPVQF..EE....DPFGLDKFLEEAKQHGGSKRPSDSSRPKEHEHEGKKRRKE~~~~~~
C.e (Q22836)      469 IIEQ..KNRFV...ADKGFSGAEGSSRG...SGPVQFEKDQ...DVFGLSSLFEHTKE...KKRGGDGGDSRGES....KRSRRD~~~~~~
A.t (gi3540201)   519 .IKN..TERFK...PDKAFTGASERV.GSKRDRPVEFEKEEEQ...DPFGLEKWVSDLKKGKKPLDKIGSGGTMRASGGGGGSSRDDDHGGSGRTK
S.p (Q09882)      502 ITSE..SRYDVLGNAHKKFKGSDEVVES..RAGPVTFEKD.IA...DPFGVDTFLNNVSSKKT~~~~~~~~~~~~~~~~~~~~~~
D.d (SNWA)        632 FGKEGGSGSGGVPRPNKEFSGTD...RSKDRTGPVAFEKEKKKSDDPFGFDDFSKKR~~~~~~~~~~~~~~~~~~~~
P.f (gi3845299)   459 PVKYIKDISDPFGL.DSLLSQAKKK~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
S.c (Prp45p)      360 HGPIQFTKAESDDKSDNYGA~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~


D.m (Bx42)        547 ~~~~~~~~~
H.s (SNW1)        537 ~~~~~~~~~
C.e (Q22836)      536 ~~~~~~~~~
A.t (gi3540201)   605 INFERSDRR
S.p (Q09882)      558 ~~~~~~~~~
D.d (SNWA)        688 ~~~~~~~~~
P.f (gi3845299)   483 ~~~~~~~~~
S.c (Prp45p)      380 ~~~~~~~~~
```

So far, the group consists of only 7 members, besides Prp45p: the 61 kDa protein Bx42 of *Drosophila melanogaster* (*D. melanogaster*, Acc.no. P39736; Frasch and Saumweber, 1989), the human 61.5 kDa protein SNW1 (U43960; Baudino *et al.*, 1998; Dahl *et al.*, 1998), a hypothetical 78 kDa protein SNWA of *Dictiostelium discoideum* (*D. discoideum*; P54705; Folk *et al.*, 1996), a hypothetical 62.7 kDa protein of *Schizosaccaromyces pombe* (*S. pombe*, Q09882), a hypothetical 69.4 kDa protein of *Arabidopsis thaliana* (*A. thaliana*; gi3540201), a hypothetical 60 kDa protein of *Caenorhabditis elegans* (*C. elegans*; Q22836) and a hypothetical 482 amino acid long protein of *Plasmodium falciparum* (*P. falciparum*; gi3845299).

Table VI.1 summarises the sequence homologies between the proteins, giving the percentage of amino acid identities of the full-length proteins, as well as of the SNW domain (underlined with a grey box in figure VI.2).

	SNW1 H.s	Bx42 D.m.	A.t.	C.e.	D.d.	S.p.	P.f.	Prp45p S.c.
SNW1 H.s		50.8 / 82.5	35.7 / 68.1	45.9 / 72.7	28.4 / 54.2	38.2 / 62.3	27.8 / 46.7	22.8 / 35.5
Bx42 D.m.	50.8 / 82.5		43.6 / 65.1	52.7 / 66.3	32.5 / 53.6	42.7 / 61.9	29.9 / 49.8	22.4 / 34.4
A.t.	35.7 / 68.1	43.6 / 65.1		40.3 / 60.2	35.0 / 54.2	39.3 / 60.2	29.7 / 51.1	23.3 / 40.0
C.e.	45.9 / 72.7	52.7 / 66.3	40.3 / 60.2		31.0 / 52.7	39.1 / 58.5	28.2 / 41.6	22.8 / 34.5
D.d.	28.4 / 54.2	32.5 / 53.6	35.0 / 54.2	31.0 / 52.7		35.4 / 53.6	23.7 / 37.0	18.6 / 30.3
S.p.	38.2 / 62.3	42.7 / 61.9	39.3 / 60.2	39.1 / 58.5	35.4 / 53.6		31.3 / 49.6	21.9 / 35.9
P.f.	27.8 / 46.7	29.9 / 49.8	29.7 / 51.1	28.2 / 41.6	23.7 / 37.0	31.3 / 49.6		24.5 / 35.4
Prp45p S.c.	22.8 / 35.5	22.4 / 34.4	23.3 / 40.0	22.8 / 34.5	18.6 / 30.3	21.9 / 35.9	24.5 / 35.4	

Table VI.1:
Homologies between the "SNW" proteins.
Pairwise sequence alignments were performed using the Genestream align program (section II.4). Values give the percentage of amino acid identity of either the full length proteins (number above line) or of the SNW domains (number below line) as defined by the grey bar in figure VI.2. Abbreviations for organisms used as in figure VI.2.

From these comparisons the degree of relationship among the SNW proteins of the different species can be deduced. The strongest relationship with respect to their primary structure is between the *Drosophila* and *Caenorhabditis* proteins (52.7% identity over the whole sequence). Although both proteins have slightly less homology to the human SNW1 protein than to each other (50.8% and 45.9%, respectively), the amino acid sequence of the SNW-domain shows a higher degree of conservation between the human and the *Drosophila* Bx42 protein (82.5%) and between the human and the *C. elegans* protein (72.7%) than it does between the *Drosophila* and the *C. elegans* protein (66.3%). This shows the close relationship of the *three* proteins and suggests an important function of the SNW domain, which remained more conserved during evolution than other parts of the three proteins.

The *S. pombe* and *Arabidopsis* proteins seem to be quite closely related. They both exhibit the highest homology to the Bx42 protein (42.7% and 43.6%), but are more distantly related to it than are the human and the *C. elegans* proteins. The *Dictostelium* protein has a 35% amino acid identity with the *S. pombe* and *A. thaliana* proteins, but shows less conservation with the other proteins. The SNW protein of *P. falciparum* exhibits approximately 30% amino acid identity to the human, the *Drosophila*, the *Caenorhabditis*, the *S. pombe* and the *Arabidopsis* protein and about 24% to the *Dictiostelium* protein and Prp45p. It is therefore difficult to put it closer to one or the other of the proteins.

The 24.5% identity to the *P. palcifarum* protein is indeed the highest score of amino acid identity of Prp45p to any of the SNW proteins. With the other SNW proteins, Prp45p shares slightly more than 20% identical amino acid residues, apart from the *Dictiostelium* protein, where the score is only 18.6%. For the SNW domain, the identities of Prp45p to the other SNW proteins range from 30-40%.

From the high degree of primary structure conservation it could well be imagined that the human, the *Drosophila*, the *S. pombe*, the *Arabidopsis*, the *Dictiostelium* and maybe the *Plasmodium* proteins represent functional homologues. Prp45p, however, seems to be more distantly related. Generally, the similarity to the other proteins is relatively low outside the SNW domain. Prp45p is considerably shorter than the other polypeptides and lacks an n-terminal extension which, in the other proteins, contains a number of very strongly conserved residues and regions (see figure VI.2).

179

Furthermore, Prp45p, in comparison to its counterparts in other organisms, lacks a conserved proline-rich region, which constitutes a potential binding site for SH3 domains (see dashed box in figure VI.2, e.g. Mongiovi *et al.,* 1999). SH3 domains (src homology domain-3) constitute a family of protein-protein interaction modules that bind to peptides displaying an X-proline-X-X-proline consensus sequence. Considering the relatively low degree of sequence conservation, it seemed uncertain whether Prp45p is in fact a functional homologue of the other SNW proteins. To clarify this question it was attempted to complement the effects of Prp45p depletion from the yeast cells by producing the human homologue SNW1 in the cells. This experiment is presented in section VI.9.

Apart from the strongly conserved SNW domain, the proteins contain a region further to the carboxy-terminus, which resembles the n-terminal half, from helix A to β-strand D), of the consensus sequence for the *src* homology domain-2 (SH2 domain; Waksman *et al.,* 1992). SH2 domains are most often found in tyrosine kinases and play an important role by recognising specifically phosphotyrosines in signal transduction cascades. They complement the action of the catalytic (tyrosine) kinase activity by communicating the phosphorylation states of signal transduction proteins to elements of the signalling pathway. Signal transduction proteins, which contain SH2 domains, but are not tyrosine kinases have also been found (for a review on src tyrosine kinases and SH2 domain function see e.g. Courtneidge *et al.,* 1993). However, it remains to be determined whether the sequence conservation of part of an SH2 domain in the SNW proteins projects to a similar function. This is particularly questionable for Prp45p, in which this domain is rather degenerate compared to other members of the SNW proteins (see figure VI.2).

To date, the functional analysis of the SNW proteins is still rudimentary. Apart from Prp45p, which will be discussed in detail later, the only members of the group for which experimental data are available are the Bx42 protein of *D. melanogaster* and the human SNW1 protein.

The Bx42 protein was identified as a nuclear protein that is tightly associated with nucleosomal chromatin (Frasch and Saumweber, 1989). It is widely expressed in

all tissues of the fly and was found to be localised to specific condensed sites on polytene chromosomes. Interestingly, this pattern of chromosomal localisation changes in response to the steroid hormone 20-OH-ecdysone (Wieland *et al.*, 1992), suggesting that Bx42 may be a steroid hormone regulated transcription factor in *Drosophila*.

For the human homologue, two groups independently also suggested an involvement of the protein in transcription. Baudino *et al.* (1998) isolated the SNW1 protein (they named it NCoA-62) in a two-hybrid screen as interactor of the nuclear vitamin D receptor (VDR). VDR acts in a heterodimeric complex with the retinoid X receptor (RXR), which binds to vitamin D-responsive promotor elements to regulate the transcription of specific genes or gene networks. The authors were able to confirm a direct interaction of SNW1 with the VDR ligand-binding domain in *in vitro* immunoprecipitation experiments using the recombinant proteins. Furthermore they demonstrated a dependence of the expression of retinoic acid-, estrogen- and glucocorticoid responsive reporter genes on co-expression of SNW1, whereas the basal level of transcription and transcription of unrelated genes was not effected. Another group isolated SNW1 (they called it Skip) as interactor of the avian retroviral oncogene v-Ski, which is also postulated to act as a transcription factor (Dahl *et al.*, 1998). SNW1 also interacted with the cellular form of the Ski protein in the two-hybrid system. The authors confirmed the observed interaction in *in vitro* experiments and demonstrated a nuclear localisation of SNW1. An involvement of v-Ski in nuclear hormone receptor pathways can well be imagined, since it was shown that v-Ski could replace the nuclear hormone receptor family oncogene v-ErbA, which transforms avian erythroid cells (Larson *et al.*, 1992).

Surprisingly with respect to the above data yielded for SNW1, the protein was shown to be a component tightly associated with the spliceosome: Neubauer *et al.* (1998) identified the protein when they assembled human splicing complexes from HeLa nuclear extracts on a biotinylated, radiolabelled pre-mRNA and purified the complexes by gel filtration and affinity chromatography. The proteins were separated by two-dimensional gel electrophoresis, the spots excised from the gel and analysed by mass spectrometry. Interestingly, three distinct spots in the gel corresponded to the SNW1 protein, suggesting that the protein might be subjected to phosphorylation or

other posttranslational modifications. The authors report furthermore, that a green fluorescent protein-SNW1 (GFP::SNW1) fusion protein colocalises with spliceosomal snRNPs.

The fact that Prp45p was isolated in the two-hybrid screen with the well characterised splicing factor Prp22p and the finding that its potential human homologue SNW1 (Skip, NCoA-62) was demonstrated to be component of the spliceosome, strongly suggests a role of Prp45p, SNW1 and possibly the other SNW proteins in the splicing process.
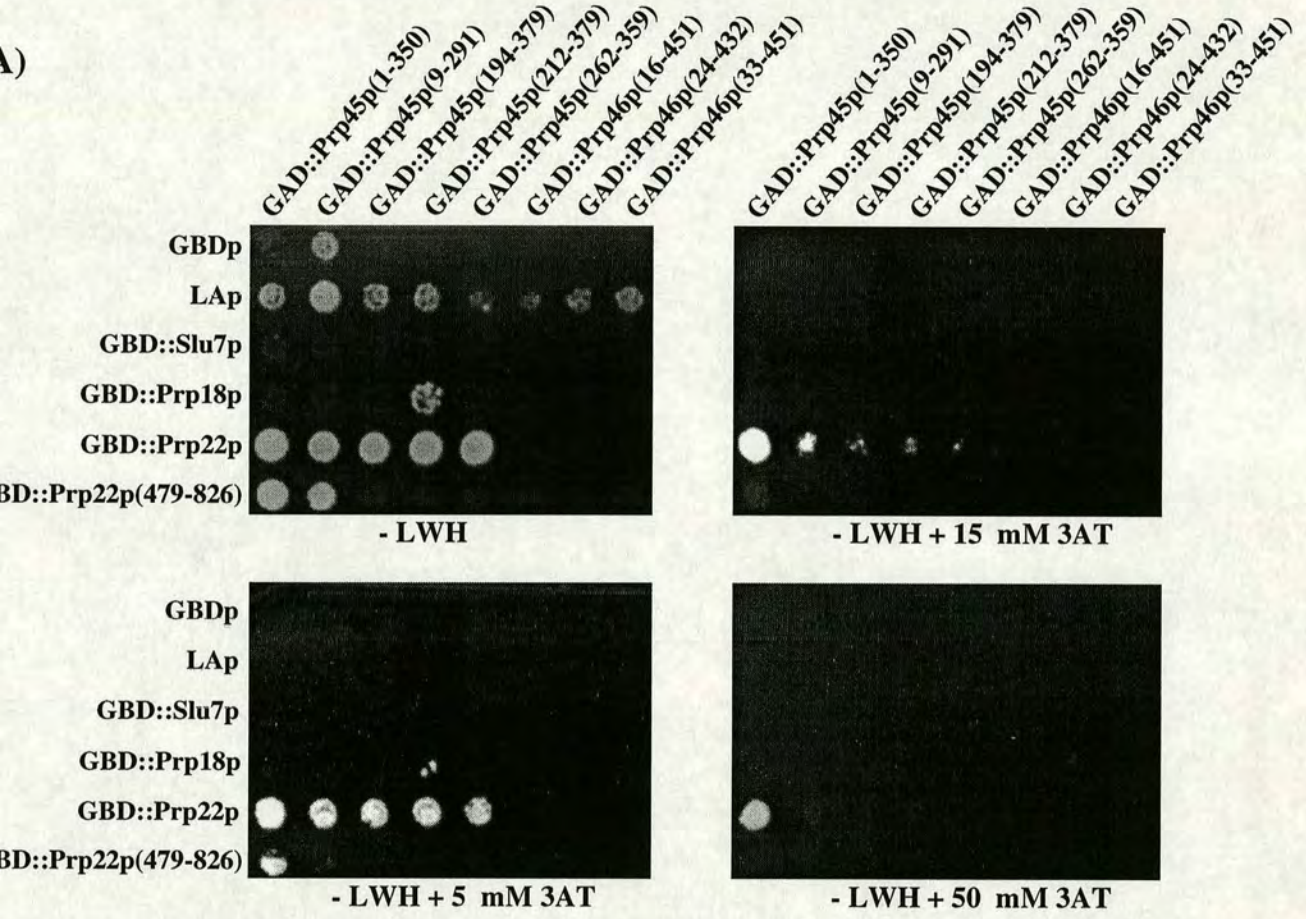
VI.3 Prp45p interacts with Prp22p in the two-hybrid system

The Prp45 protein was identified in a two-hybrid screen using Prp22p as bait (see chapter V.3). Five different prey plasmids were isolated that encoded fusion proteins containing different fragments of Prp45p (see also figure V.1 A, Chapter V.3). Fragments Prp45p(1-350) and Prp45p(9-290) cover large parts of the n-terminus and stretch into the c-terminus of the 379 amino acids long protein. These fragments are statistically most significant, because they were identified in the screen 5 and 2 times, respectively. Three other fragments Prp45p(212-379), Prp45p(262-359) and Prp45p(194-379), that were each found only once, cover large parts of the c-terminus, but do not contain n-terminal regions. The common domain of all prey fragments, comprises amino acids 262-291. The strength of the interaction of these fragments with Prp22p was investigated in more detail in a directed two-hybrid mating assay. Figure VI.3A shows the results of this assay in which the strength of the interaction was monitored by adding different concentrations of 3AT to the medium.
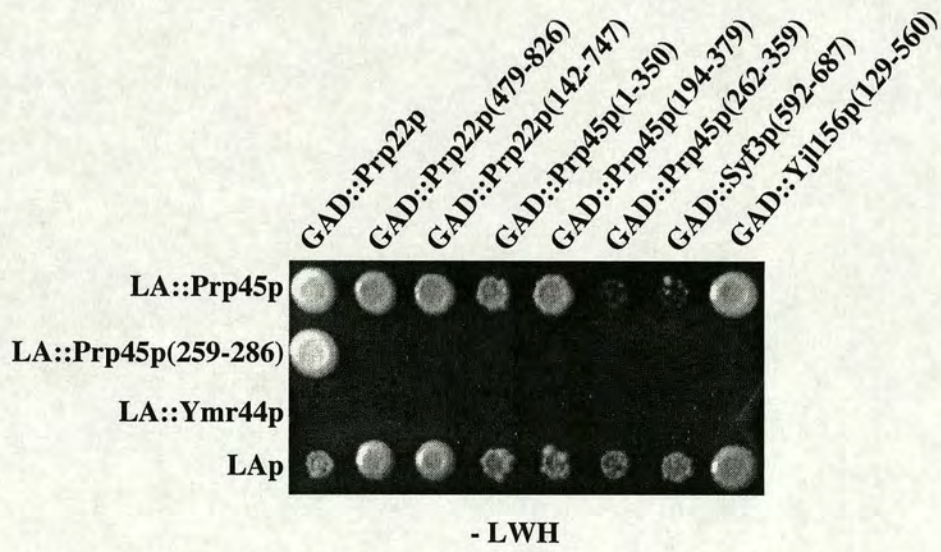
It is demonstrated that all Prp45p prey fragments support a two-hybrid interaction with Prp22p up to 3AT concentrations of at least 15 mM. Prp45p(1-350) even allows growth of the diploid cells up to a 3AT concentration of 50 mM. Furthermore it can be seen that the frequency with which the prey fragments were isolated in the screen is reflected in the strength of the two-hybrid interaction. Prp45p(1-350), found 5 times, interacts strongest with Prp22p. Prp45p(9-291), isolated twice, interacts slightly stronger than the c-terminal three fragments.

Figure VI.3: Investigation of the Prp22p-Prp45p interaction in a directed two-hybrid mating assay. CG1945 and Y187 yeast strains were transformed with pAS2ΔΔ-, pBTM116- or pACTIIStop-derived plasmids encoding the depicted bait or prey fusion proteins. Transformants were propagated on the appropriate selective medium and then mated on YPDA medium overnight at 30°C. The resulting diploids were propagated on YMM -WL medium for 24 h at 30°C and transferred onto YMM -LWH selective medium containing different concentrations of 3AT. Cells were incubated for 2 days at 30°C. **A)** Bait plasmids used were: GBDp (Gal4p DNA-binding domain), pASΔΔ; LAp (LexA protein), pBTM116; GBD::Slu7p, pASSlu7; GBD::Prp18p, pASPrp18; GBD::Prp22p, pMA22; GBD::Prp22p(479-826), pMA22s. Prey plasmids used were: GAD::Prp45p (1-350), p22-1; GAD::Prp45p(9-291), p22-10, GAD::Prp45p(194-379), p22-13; GAD::Prp45p(212-379), p22-19; GAD::Prp45p(262-359), p22-27; GAD::Prp46p (16-451), p45-31; GAD::Prp46p(24-432), p45-3; GAD::Prp46p(33-451), p45-45; **B)** Bait plasmids used were: LA::Prp45p, pMA45; LA::Prp45p(259-286), pMA45cr; LA::Ymr44p, pMA44BTM; LAp, pBTM116; Prey plasmids used were: GAD::Prp22p, p22; GAD::Prp22p(479-826), p7-III-5; GAD::Prp22p(142-747), p7-III-9; GAD::Prp45p(1-350), p22-1; GAD::Prp45p(194-379), p22-13; GAD::Prp45p(262-359), p22-27; GAD::Syf3p(592-687), p22-12; GAD::Yjl156p (129-560), p7-IV-27; (GAD - Gal4p transcription activation domain).

A)

	GAD::Prp45p(1-350)	GAD::Prp45p(9-291)	GAD::Prp45p(194-379)	GAD::Prp45p(212-379)	GAD::Prp45p(262-359)	GAD::Prp46p(16-451)	GAD::Prp46p(24-432)	GAD::Prp46p(33-451)
GBDp								
LAp								
GBD::Slu7p								
GBD::Prp18p								
GBD::Prp22p								
BD::Prp22p(479-826)								

- LWH

- LWH + 15 mM 3AT

- LWH + 5 mM 3AT

- LWH + 50 mM 3AT

B)

	GAD::Prp22p	GAD::Prp22p(479-826)	GAD::Prp22p(142-747)	GAD::Prp45p(1-350)	GAD::Prp45p(194-379)	GAD::Prp45p(262-359)	GAD::Syf3p(592-687)	GAD::Yjl156p(129-560)
LA::Prp45p								
LA::Prp45p(259-286)								
LA::Ymr44p								
LAp								

- LWH

The insert encoding Prp22p(479-826), which was originally found as prey fragment in the Slu7p two-hybrid screen (see figure III.2, chapter III.2.2), was cloned into the pASΔΔBC bait vector to produce the fusion protein GBD::Prp22p(479-826). This central fragment of Prp22p contains all 6 conserved motifs of the DEAH-box family. It is shown here, that this fragment of Prp22p is able to interact with Prp45p, at least with prey-fragment Prp45p(1-350). However, the interaction of Prp45p(1-350) with this Prp22p fragment is not as strong as the interaction with the full length Prp22p bait, suggesting that although Prp45p interacts with Prp22p in the conserved region, the interaction is probably strengthened by contacts between the proteins outside this highly conserved region of Prp22p.

Prp45p was also tested in this experiment for an interaction with Slu7p and Prp18p, which are known to act at least at one point of the splicing process in close proximity to Prp22p. None of the Prp45p prey fragments interacts with Slu7p. They also do not interact with Prp18p, except fragment Prp45p(212-379), for which a two-hybrid interaction with Prp18p allows weak cell growth on a 3AT concentration of up to 5 mM. However, since Prp45p(212-379) differs from fragment Prp45p(194-379) only in that it contains 19 aminoacid residues less at its n-terminus, it is questionable whether the observed interaction is biologically important or an artefact of the system. Maybe, just in this particular fusion, the Prp45p fragment folds into a certain way distinct from the native structure, which is prone to form an aspecific interaction with the Prp18 protein. Fusion proteins of Prp46p (GAD::Prp46p(16-451), GAD::Prp46p(24-432) and GAD::Prp46p(33-451), (see section VI.10.2, fig. VI.16) were used here as a negative control, in order to show that the Prp22p bait does not interact with any prey tested.

The distribution of the Prp45p prey fragments found with Prp22p, a group of more amino-terminal fragments (Prp45p(1-350) and Prp45p(9-291)) and a group of carboxy-terminal fragments (Prp45p(212-379), Prp45p(262-359) and Prp45p(194-379)), that share a small overlap, leaves mainly two possibilities for the interaction of Prp45p with Prp22p. Either there are separate contact points of Prp22p with n- and c-terminal regions of Prp45p, or the 30 amino acid overlap between the prey-fragments is itself crucial for the protein-protein interaction.

To address this question it was investigated whether the common domain of the prey fragments alone was sufficient to support a two-hybrid interaction of Prp22p with Prp45p. For this reason, oligonucleotides FCR-1 and FCR-2 were used in a PCR to amplify the region encoding amino acids 259-286 of Prp45p from plasmid p22-1 (encoding fragment GAD::Prp45p(1-350)). The region deviates slightly from the actual common region of amino acids 262-291, since primer sequences were chosen that guaranteed optimal annealing to the template DNA. The amplified and gel-purified fragment was subsequently cloned into the polylinker of bait vector pBTM116 via the *Bam*H1 restriction site. The resulting plasmid was named pMA45cr and should express a fusion protein of Prp45p(259-286) fused to the LexA protein (LA::Prp45p(259-286)). The integrity of the insert (orientation and sequence) was checked by DNA sequencing. The plasmid was then transformed into yeast strain L40. A directed two-hybrid mating assay was performed to test the small fusion protein LA::Prp45p(259-286) for an interaction with Prp22p preys (figure VI.3B). Full-length Prp45p fused to the LexA protein (LA::Prp45p) produced from plasmid pMA45 was also tested (the cloning of pMA45 will be described in section VI.10.1).

The LexAp (LAp) as well as the LA::Prp45p has a potential to autoactivate transcription when produced together with a number of prey fusions in the diploid cells (bottom and top lane). For that reason the results yielded for the full length Prp45p fused to LAp are difficult to interpret (top lane). Although cell growth seems to be considerably increased when the proteins LA::Prp45p and GAD::Prp22p are co-expressed, compared to cell growth upon co-expression of LAp and GAD::Prp22p, it cannot be stated clearly whether this weak difference can be accounted to a strengthened two-hybrid interaction or simply to different expression levels or differences in stability of the proteins. However, if the common region of Prp45p is fused to the LAp (LA::Prp45p(259-286), the activation of the reporter genes is more specific and an interaction can only be detected with the full length Prp22p prey (GAD::Prp22p). This suggests that indeed the region spanning the amino acids 259-286 of Prp45p is important and sufficient for an interaction with Prp22p. The interaction is considerably weaker, though, than the interaction of the different Prp45p prey fragments with the Prp22p bait shown above (figure VI.3A). On plates containing 5mM 3AT, cell growth is abolished, as is true of all interactions shown in figure VI.3B

(data not shown). This indicates that additional amino acid residues in Prp45p might be involved in supporting the interaction or affecting the production or stability of the fusion protein. These residues are likely to be located further to the n-terminus of the protein, as is suggested by the strong interactions of the prey fragments Prp45p(1-350) and Prp45p(9-291) with Prp22p.

Interestingly, two putative phosphorylation sites are located within this Prp22p-interacting region of Prp45p (see also figure V.1A, chapter V.3). A threonine residue at position 267 and a tyrosine residue at position 278 are potential targets for protein kinases, since the preceding amino acid residues constitute recognition motifs for serine/threonine and tyrosine kinases, respectively.

VI.4 Generation of a conditionally regulated and protein A-tagged *PRP45* allele

In order to investigate a potential function of Prp45p in nuclear pre-mRNA splicing, a chromosomal *PRP45* allele was generated, the expression of which could be conditionally regulated. This should allow the depletion of the protein from the cells under non-inducing conditions and a subsequent investigation of the consequences of this depletion for the splicing process. At the same time the protein should be tagged, providing a tool to immunoprecipitate it from cell extracts in order to study a possible association with spliceosomal components.

To achieve this aim a one-step PCR-mediated strategy for the construction of conditionally expressed genes developed by Lafontaine and Tollervey (1996) was applied. Using oligonucleotide primers PFun-1 and PFun-2, a cassette was PCR-amplified from plasmid pTL27 containing the *HIS3* marker gene, the yeast *Gal1-10* promotor region and a sequence that encodes two IgG binding domains of the *Staphylococcus aureus* protein A. The PCR product was separated from the vector by agarose gel electrophoresis and purified using the QIAquick Gel Extraction Kit (Qiagen). Subsequently, the cassette was transformed into the diploid yeast strain BMA38. The flanking sequences of the primers were chosen to allow homologous

recombination of the cassette upstream of the chromosomal *PRP45* gene, in such way that the protein A sequence becomes fused to the ATG start codon of the *PRP45* ORF and 100 bp of the *PRP45* promotor region directly upstream the ATG is deleted. Thus, integration of the cassette at the desired locus should put the *PRP45* ORF under control of the *Gal1* promotor, i.e. one should be able to allow expression of this *PRP45* allele by growing the strain with galactose as sole carbon source and shut off its expression by growing the strain on glucose-based medium. Transformants were selected by plating the cells onto medium lacking histidine and containing 2% galactose, 2% raffinose and 2% sucrose (YMGRS -H) but no glucose. 140 histidine prototrophic transformants were obtained.

The integration of the cassette into the chromosomal locus was checked by PCR on one of the yeast colonies using oligonucleotides PFun1 and PFun2. A DNA fragment of the expected size of approximately 3 kb could be amplified from this transformant but not from the wt parental strain, giving a first indication that the cassette had been integrated into the chromosome at the desired location in this clone (data not shown). This yeast clone was named YMA45/1.

Subsequently a western analysis was performed to look at the production of the protein A-tagged Prp45p in this strain, to prove that the 3 kb PCR product was indeed derived from the cassette integrated in the promotor region of the *PRP45* gene and to demonstrate that *PRP45* gene expression can be shut off by growing the strain under non-inducing conditions. The YMA45/1 strain as well as the parental wt strain BMA38 were grown for 15 hours in either YPDA or YMGRSsup liquid medium and subsequently crude cell extracts were prepared from the cultures. Aliquots of the cell extracts were then run on a 12% SDS-polyacrylamide gel and electroblotted. The protein A-tagged Prp45p was detected using anti-rabbit IgG-horseradish peroxidase linked antibodies and visualised by ECL (see figure VI.4).
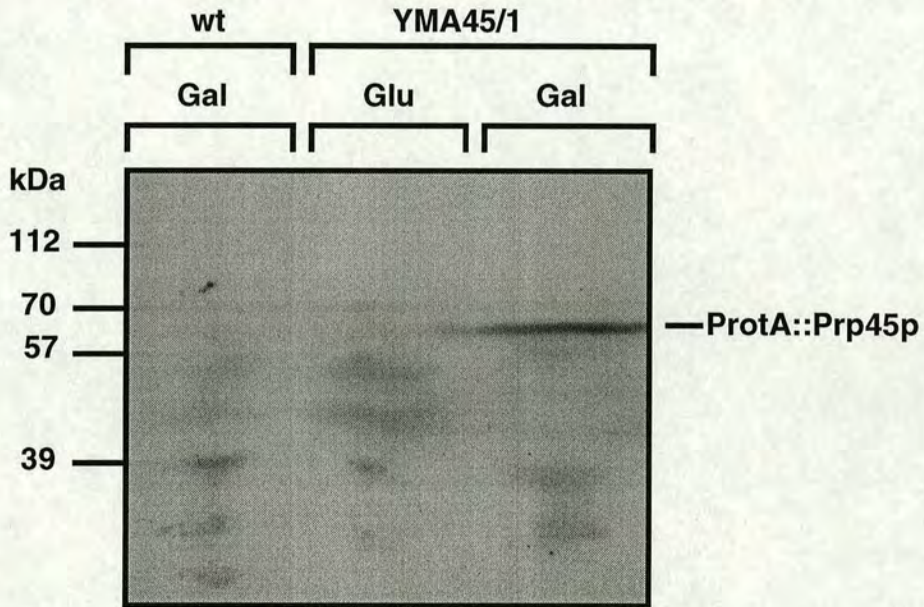
A)
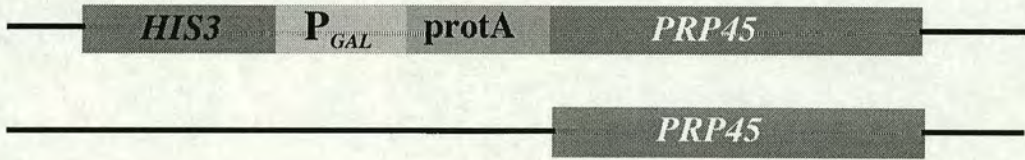
B)

Figure VI.4:
Western blot analysis of protein A-tagged Prp45p.
A) Schematic illustration of the expected integration of the *HIS3-P$_{GAL}$*-protA regulation cassette into the genomic *PRP45* locus in strain YMA45/1. **B)** Western blot analysis to investigate the regulative production of the protein A-tagged Prp45p fusion protein (ProtA::Prp45p). Strain YMA45/1 was grown for 15 hours at 30°C in either YPDA (Glu) or YMGRSsup (Gal) liquid medium and crude cell extracts were prepared (wt strain BMA38 grown in YMGRSsup was used as control). The proteins were fractionated on a 12% (w/v) SDS-polyacrylamide gel, electroblotted and probed with anti-rabbit IgG-horse radish peroxidase antibodies. Proteins were visualised by ECL. The position of the ProtA::Prp45p is indicated.

It can be seen that under inducing conditions (growth in YMGRSsup) a protein which runs at about 65 kDa is produced in cells from strain YMA45/1. This protein is absent when cells of that strain were grown in YPDA . This strongly suggests, that indeed the cassette had been integrated into the *PRP45* promotor region, thereby putting the *PRP45* gene under control of the *Gal1* promotor. Under inducing conditions, the protein A-tagged Prp45p can be detected as a clear single band. Glucose repression of the *Gal1* promotor, however, prevents the production of the tagged protein, at least to an extent that makes it undetectable under the conditions used in this experiment. Protein A-tagged Prp45p runs slightly slower than expected. Its calculated molecular weight is 57.2 kDa (42.5 kDa for Prp45p + 14.7 kDa for the double protein A epitope). However, Prp45p is a highly charged and basic protein, which could account for its slightly abnormal electrophoretic mobility.

In order to investigate the effects of the depletion of Prp45p on cell growth as well as on pre-mRNA splicing, a haploid strain containing the conditionally regulated *PRP45* gene had to be generated. For this reason the diploid strain YMA45/1 was sporulated and the resulting tetrads were dissected onto YMGRSsup solid medium as well as onto YPD agar plates (figure VI.5).

As expected, on YPD plates in any case at most two spores germinated. Obviously, depletion of Prp45p prevents spore germination and/or cell growth. On galactose-based solid medium, under conditions that should allow the production of the tagged protein, in half of the dissected tetrads 3 or 4 spores germinated and grew to colonies, indicating that the production of the tagged protein is able to support cell growth. The cells grown from spores of tetrad 3 (figure VI.5A) were subsequently suspended in microtiterplates and spotted onto either YPDA- or YMGRSsup agar plates (figure VI.5B). As expected, only in two cases the cells grew to colonies on YPDA agar, whereas on YMGRSsup agar plates all cells formed colonies. Two spores should contain the regulated *PRP45* allele. On YPDA solid medium, the essential Prp45 protein becomes depleted, therefore two spores do not form colonies. The fact that on YMGRSsup plates all colonies grew essentially equally fast at a range of temperatures (16, 23, 30 and 37°C) indicated that tagging Prp45p at its n-terminus does not effect its function, at least not to a significant extent in this assay.

190

Figure VI.5:
Effects of metabolic depletion of Prp45p.
A) Tetrad analysis of strain YMA45/1. Strain YMA44/1 was grown overnight at 30°C in YMGRSsup, the cells collected and transferred to Spm sporulation medium. After 4 days of incubation at 23°C the formed tetrads were dissected onto either YPD (Glucose) or YMGRSsup (Galactose) agar plates. The spores were incubated for four days at 23°C. **B**) Haploid cells grown from the spores of tetrad 3 (Galactose, figure A)) were serially diluted in microtiter plates and spotted onto either YMGRSsup (Galactose) or YPDA (Glucose) agar plates and incubated at a range of temperatures for 2 days.

A)

Glucose | Galactose

B)

Galactose | Glucose

37 °C

30 °C

23 °C

16 °C
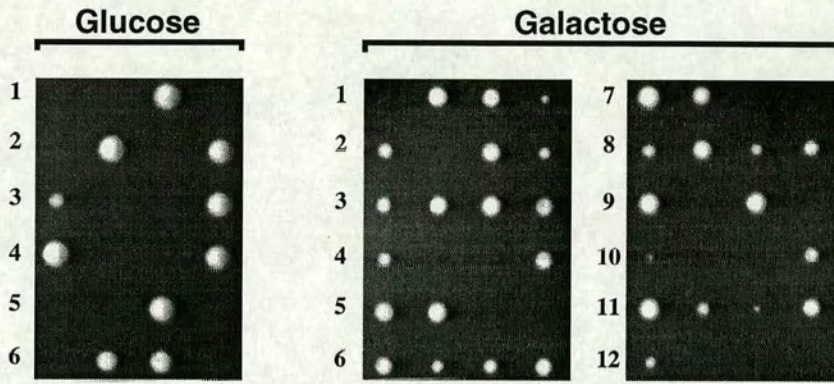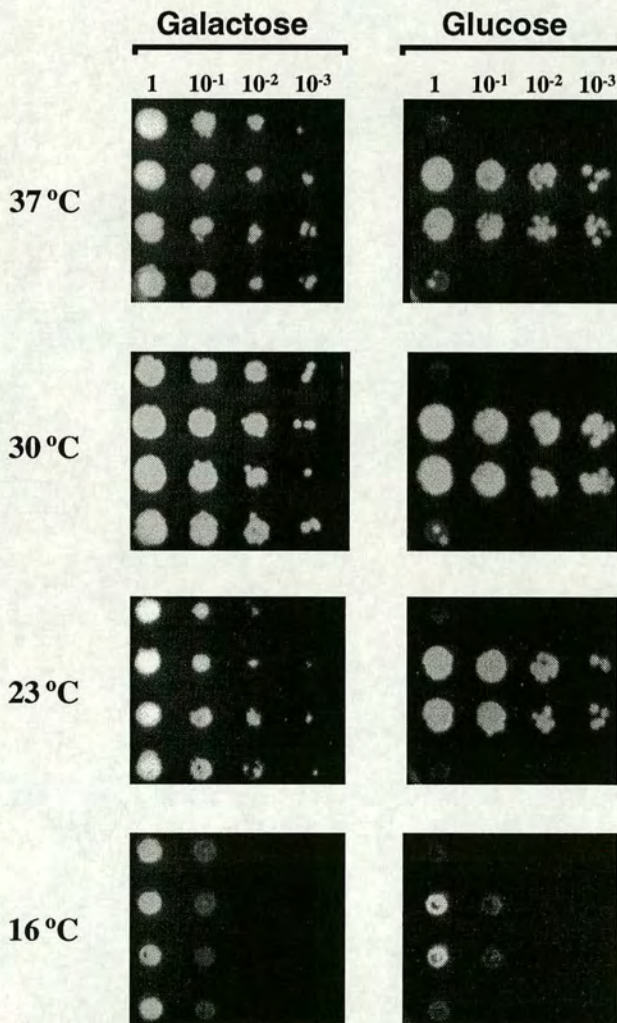
By restreaking all colonies depicted in figure VI.5B onto YMGRS -H solid agar plates, it could be confirmed, that indeed the cells that did not grow on glucose were histidine prototrophs. This demonstrated that the growth defect on glucose segregated with the *HIS3* marker gene and therefore is due to the insertion of the cassette into the chromosome (data not shown). The haploid strain containing the conditionally regulated *PRP45* gene that encodes the tagged protein will be subsequently referred to as YMA45/2.

To study the effect of Prp45p depletion on cell growth in more detail, the growth of the strains YMA45/2 and BMA38n was compared. Both strains were grown in YPDA- and YMGRSsup liquid medium at $30^{\circ}C$ and growth was monitored by taking samples at different time points to measure the optical density at 600 nm wavelength (OD_{600}). In figure VI.6A it is confirmed, what was already demonstrated in figure VI.5B for growth on solid medium, namely that there is no difference in cell growth of strain YMA45/2 and the wt strain BMA38n under permissive conditions. This means that neither changing the expression level of Prp45p (*PRP45* is likely to be overexpressed under the strong *Gal1* promotor), nor adding an n-terminal protein A tag effects the function of the essential protein to an extent that would be visible in these growth assays. Figure VI.6B shows the effect of Prp45p depletion on cell growth. After about 7 to 8 hours growth in YPDA liquid medium, growth of strain YMA45/2 starts to slow down compared to the wt strain and after about 13 hours the cells stop doubling, and growth levels off. The data of the Prp45p depletion experiments confirm the finding of Diehl and Pringle (1991), who demonstrated by disruption of the *PRP45* ORF that *FUN20* (*PRP45*) is an essential gene.

Figure VI.6:
Effect of Prp45p-depletion on cell growth.
Overnight cultures of strains YMA45/2 and BMA38n were grown in YMGRSsup liquid medium to midlogarithmic phase and aliquots of the cultures were used to inoculate pre-warmed 250 ml of either YMGRSsup- or YPDA liquid medium to an OD_{600} of approximately 0.1. The cultures were grown at $30^{O}C$ and aliquots were removed in order to monitor growth by measuring the OD_{600} at different time points. The cultures were diluted at intervals to maintain logarithmic growth. **A)** Growth of BMA38n and YMA45/2 in YMGRSsup (Galactose-based) liquid medium. **B)** Growth of BMA38n and YMA45/2 in YPDA (Glucose-based) liquid medium.
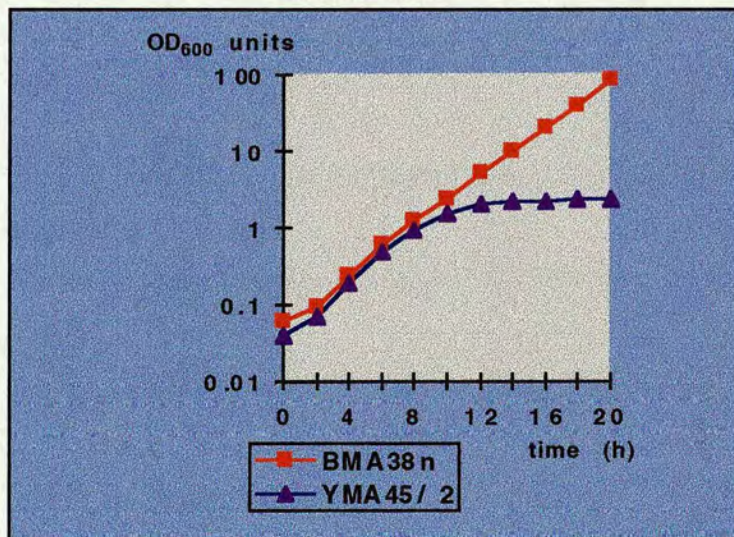
A)

Galactose:
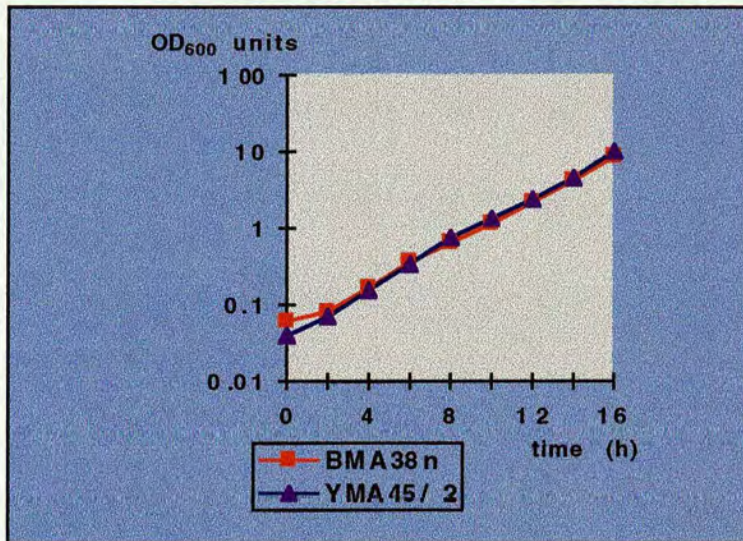
B)

Glucose:

In order to finally prove that the observed detrimental effect on cell growth in glucose-based medium is in fact solely due to the depletion of Prp45p, complementation of the growth defect of strain YMA45/2 by providing Prp45p was tested (figure VI.7).
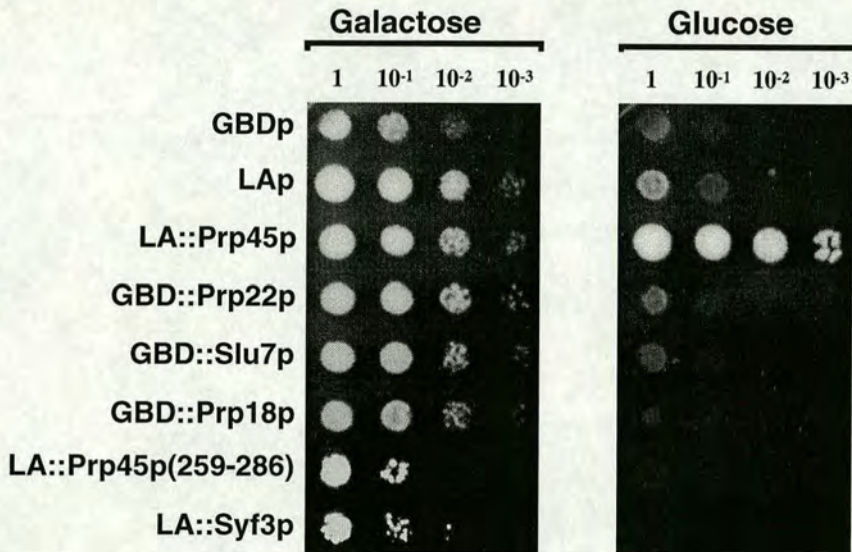

Figure VI.7:
Complementation of Prp45p depletion by the LexA::Prp45 fusion protein. pASΔΔ- or pBTM116-derived bait plasmids encoding the depicted fusion proteins were transformed into the YMA45/2 strain. Transformants were selected on YMGRS -W medium, restreaked for colony purification and then suspended, serially diluted, into microtiterplates and spotted onto either YMGRS -W (Galactose) agar or YMM -W (Glucose) agar. Cells were incubated for 3 days. Bait plasmids used were: GBDp (Gal4p DNA-binding domain), pASΔΔ; LAp (LexA protein), pBTM116; GBD::Slu7p, pASSlu7; GBD::Prp18p, pASPrp18; GBD::Prp22p, pMA22; LA::Prp45p, pMA45; LA::Prp45p(259-286), pMA45cr; LA::Syf3p, pCR117.

Indeed, transformation of YMA45/2 with bait plasmid pMA45, from which a LexA::Prp45p (LA::Prp45p) fusion protein is produced (section VI.10.1) complements the growth defect, whereas producing only LexAp (LAp) or other fusion proteins does not. Notably, overexpression of the second step splicing factors Slu7p, Prp18p or Prp22p does not overcome the effect of Prp45p depletion. This suggests that Prp45p

196

does not simply support the function of these proteins by, for example, facilitating binding of these factors to the spliceosome, in which case overexpression of these factors might compensate for the loss of Prp45p function. Not surprisingly, the small fragment of Prp45p (Prp45p(259-286)), which was found to interact with Prp22p in the two-hybrid assay, is also not able to complement the growth defect of strain YMA45/2, suggesting that Prp45p has other essential functions, besides the binding of Prp22p. These data prove that in fact the depletion of Prp45p is responsible for the growth defect of strain YMA45/2 on YPDA solid- and in YPDA liquid medium. Furthermore, they demonstrate the functionality of the Prp45p bait fusion protein, which was later used in a two-hybrid screen to search for Prp45p-interacting proteins (section VI.10.2).

In addition, the above experiments illustrate that the regulated expression system works, i.e. that *PRP45* gene expression can be effectively turned off by growing the cells of strain YMA45/2 in glucose-based medium. Therefore, the strain can be used to investigate whether the depletion of Prp45p has an effect on pre-mRNA splicing.

VI.5 Prp45p is required for pre-mRNA splicing *in vivo*

The strong two-hybrid interaction of Prp45p with Prp22p suggested a role for Prp45p in pre-mRNA splicing. Not only the strength of the interaction, but also its specificity was intriguing. Prp45p had not been identified in any previously performed two-hybrid screens within the TAPIR network, indicating that the protein does not contain any structural features or motifs, which would give rise to aspecific interactions with other proteins or RNAs that could provoke false positive two-hybrid responses.

The generation of strain YMA45/2, in which the *PRP45* gene can be conditionally expressed, now allowed the initial functional characterisation of Prp45p. To directly investigate a potential role of Prp45p in the splicing process, the effect of Prp45p depletion on pre-mRNA processing was studied.

197

Cultures of strains YMA45/2 and BMA38n were grown in either galactose-(YMGRSsup) or glucose-(YPDA) based medium and samples were taken at different time points. Subsequently total RNA was prepared from those cells and the processing of pre-mRNA was monitored by Northern analysis (figure VI.8).

When probing against exon 1 of the *RP28* gene, a strong accumulation of pre-mRNA was clearly visible in strain YMA45/2, when cells were grown under non-permissive conditions, i.e. in YPDA liquid medium. After 8 hours growth under repressing conditions, unprocessed precursor mRNA could be seen and the amounts increased further after prolonged growth (see 10 and 12 hour time points). No pre-mRNA could be detected in cells of either the YMA45/2 strain grown under permissive conditions (on galactose-based medium) or the BMA38n wt control strain grown in YPDA medium. These data show a clear splicing defect in cells depleted of Prp45p and thus demonstrate a requirement for Prp45p for the processing of pre-mRNA. Concomitantly with the accumulation of the precursor mRNA a decrease in the amount of mature message was observed. This was most clearly seen for the 12 hours time point, in which the detected mRNA levels were decreased under non-permissive conditions compared to permissively grown cells, although the loading control (*PGK*-mRNA, bottom panel) suggested, that even less RNA was loaded in the latter case. Therefore, it could be concluded that Prp45p is a bona fide splicing factor, required for efficient pre-mRNA splicing *in vivo*.

The northern analysis confirmed, what was already suggested when the growth of the YMA45/2 strain was compared to wt growth under permissive conditions, namely that the protein A-tagged Prp45p (protA::Prp45p) is functional. Pre-mRNA splicing *in vivo* does not seem to be effected by the n-terminal tag nor by putatively increased expression levels of the *PRP45* gene under control of the strong *Gal1* promotor.
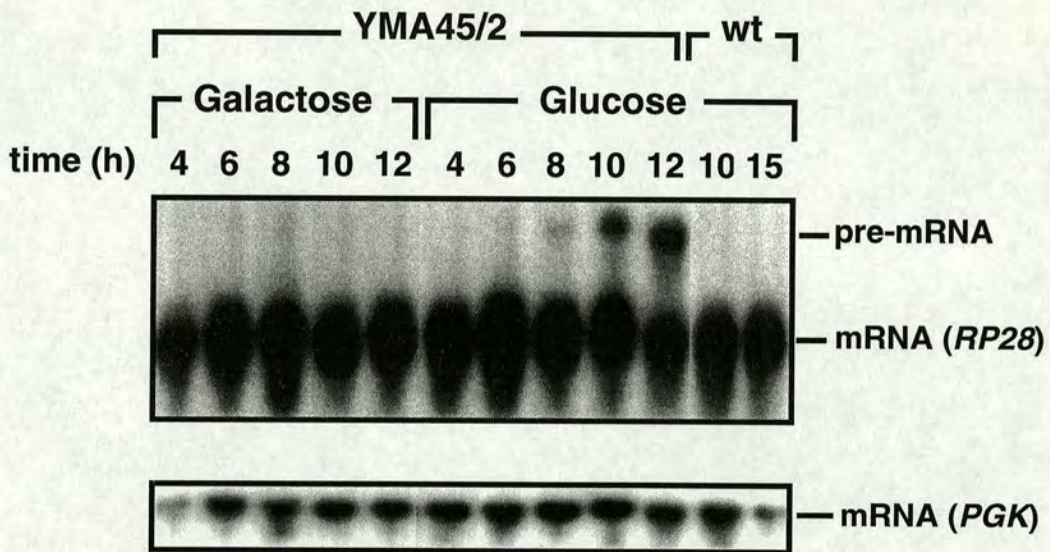
Figure VI.8:
Northern analysis of splicing upon Prp45p depletion. Cells of strain YMA45/2, as well as BMA38n wild-type cells were grown (as described in figure VI.6) in YMGRSsup (Galactose) or YPDA (Glucose) liquid medium at 30°C. 10 ml aliquots of the cultures were removed after different incubation times and total RNA was extracted from the cells. 10 μg of RNA was loaded on a 1% (w/v) formaldehyde gel and blotted to Hybond-N membrane (Amersham). The blot was then probed with a radiollabelled DNA fragment complementary to exon 1 of the *RP28* gene. The result was visualised by autoradiography. The positions of the *RP28* pre-mRNA and mRNA are indicated. The blot was stripped and reprobed with a radiolabelled DNA fragment of the intronless *PGK* gene in order to control loading (bottom panel).

Since Prp45p was found as a Prp22p interactor, the question arises whether it is a general splicing factor, required for splicing of all introns or, as for Prp22p, Slu7p and Prp18p, whether its requirement might be dictated by the distance between the branchpoint and the 3' splice site. Prp22p for instance, is only required for splicing of introns in which the branchpoint-3' splice site distance is larger than or equal to 21 nucleotides (Schwer and Gross, 1998). Therefore, if Prp45p interacted with Prp22p during 3' splice site recognition and cleavage, it would not be surprising to see a similar pattern of selective Prp45p requirement.

The *RP28* pre-mRNA, which was used in the northern analysis, contains an intron in which the branchpoint-3' splice site distance is 39 nucleotides. In order to test whether Prp45p might be dispensable for splicing of introns with very short branchpoint-3' splice site distances, splicing of U3 snoRNA was investigated. The distance between the branchpoint sequence and the 3' terminal CAG of the U3 intron is only 7 nt. Again cultures of YMA45/2 and the wt strain BMA38n were grown under permissive and restrictive conditions and samples were taken at different time points. Total RNA was isolated from these cells and the processing of U3 precursor RNA was investigated, this time by primer extension analysis. An oligonucleotide (U3exon2) hybridising to the extreme 5' end of exon 2 was used for the extension reaction, so that the removal of the U3 intron could be monitored. Extension products were expected to be 236 and 209 nucleotides long for the U3A and U3B precursor RNAs, respectively, 79 nucleotides for the mature U3 snoRNA and 78 nucleotides for the intron-lariat exon 2 intermediate. The intron-free U1 snRNA served as an internal loading control. Its extension product should be 165 nucleotides in length. The results of this primer extension assay are depicted in figure VI.9.

Figure VI.9:
Primer extension analysis of splicing upon Prp45p depletion. Cells of strain YMA45/2, as well as BMA38n wild-type cells, were grown (as described in figure VI.6) in YMGRSsup (Galactose) or YPDA (Glucose) liquid medium at 30°C. 10 ml aliquots of the cultures were taken after different incubation times and total RNA was extracted from the cells. 10 µg of RNA was used in a primer extension reaction using radiolabelled oligonucleotide primers complementary to the extreme 5' end of exon 2 of the U3 snoRNA and to the (intronless) U1 snRNA. The reactions were deproteinised, extracted with phenol/chloroform and the products resolved on a 6% (w/v) polyacrylamide gel. The result was visualised by autoradiography. The positions of the extension products are indicated.

It is clearly demonstrated here, that the depletion of Prp45p (YMA45/2 grown in glucose-based medium) leads to a strong accumulation of the two U3 precursor RNAs. Already after 6 hours growth under non-permissive conditions the two pre-U3 RNAs were strongly detectable. Only very weak bands of pre-U3A RNA appear at late time points in the control tracks (strain YMA45/2 grown in galactose-based medium and BMA38n grown in YPDA). Therefore, it could be concluded that Prp45p is required for splicing of U3 snoRNA *in vivo* and this finding strongly suggests that the requirement of Prp45p for pre-mRNA splicing is not restricted to introns with 3' splice sites distal to the branchpoint, but instead that Prp45p represents a general splicing factor, required all splicing reactions. Furthermore, this result argues against the hypothesis that Prp45p directly modulates Prp22p function in 3' splice site selection and cleavage. One would expect to see a similar selective requirement of Prp45p only for splicing of introns with distal branchpoint-3' splice site distances in this case. However, it is still possible that Prp45p acts as a general second step splicing factor in close proximity or even in contact with Prp22p at that stage of the splicing process. Equally likely would be that Prp45p interacts with Prp22p at another stage of the splicing reaction.

To address this question, the association of Prp45p with the spliceosome and with spliceosomal snRNPs was analysed, in the hope to gain knowledge about the timepoint of action of Prp45p during the processing of pre-mRNAs. These investigations will be described in the next section.

VI.6 Prp45p associates with the spliceosome

## VI.6.1 Prp45p coprecipitates spliceosomes

In order to determine whether and at what particular stage of the splicing process, Prp45p is associated with the spliceosome, it was tested whether the protein A-tagged Prp45p (protA::Prp45p) was able to coprecipitate precursor, intermediates or products of splicing. For that reason protA::Prp45p was produced in strain YMA45/2

and whole cell extracts (splicing extracts) were prepared. As control, a splicing extract was prepared from strain BMA38n expressing the double protein A epitope from vector pNOPPATA1L (kind gift of Klaus Hellmuth). Using radiolabelled actin pre-mRNA as a substrate a splicing assay was performed. After 30 min, 10 % of the sample was analysed for splicing activity, while the remaining 90 % was incubated with immunoglobulin agarose (IgG-agarose) or as controls, agarose beads or anti-Prp8p antibodies. After the immunoprecipitation, the precipitates were treated with proteinase K, extracted with phenol/chloroform and analysed for the presence of radiolabelled RNA. The samples incubated with agarose beads (no antibodies) served as negative control to determine the background level of RNA precipitation due to aspecific binding of RNA to the agarose beads. The extract in which the protein A epitope (protA) was produced represented another negative control to establish the specificity of RNA precipitation by the Prp45p protein. Using anti-Prp8p antibodies served as a positive control. Prp8p is a U5 snRNP protein which is associated with the spliceosomes throughout the splicing reactions and should therefore precipitate pre-mRNA, intermediates and the products of the splicing reaction as shown previously by Teigelkamp et al. (1995).

To determine in more detail at which timepoint Prp45p might associate with the spliceosome, the coimmunoprecipitations with protA::Prp45p were also performed when splicing was blocked at a particular step. This was achieved by adding purified recombinant proteins of dominant negative (dn) alleles of the splicing factors Prp2p (dn Prp2^{LAT}p, aminoacid exchange S378L) and Prp16p (dn Prp16^{AKT}p, aminoacid exchange G378A) to the extracts before the splicing reaction was performed. Prp2p interacts transiently with the spliceosome immediately prior to the first transesterification reaction (Kim and Lin, 1993; King and Beggs, 1990). The dominant negative Prp2^{LAT}p binds to the spliceosome but fails to function correctly. The protein cannot be released from the spliceosome and splicing becomes stalled prior to the first step (Plumpton et al. 1994). The dn Prp16^{AKT}p behaves in a similar way, but stalls splicing after the first but prior to the second step. The recombinant dn Prp16^{AKT}p was generated and purified by Ian Dix and Caroline Russell, dn Prp2^{LAT}p was a generous gift from Margaret McGarvey (all this laboratory). In figure VI.10 the results of the coimmunoprecipitation experiments are presented.

Figure VI.10:
Coprecipitation of spliceosomes by Prp45p. Whole yeast cell extract (splicing extract) was prepared from cells of strain YMA45/2, grown in YMGRSsup liquid medium, producing the protein A tagged Prp45p (protA::Prp45p). As a control, an extract was prepared from strain BMA38n transformed with vector pNOPPATA1L, producing a double protein A epitope (protA). Splicing was performed in the extracts by addition of radiolabelled actin precursor-mRNA and splicing buffer (50 μl total volume). The reactions were stopped and 5 μl was removed (splicing controls, input), deproteinised and the products kept at -70°C for precipitation until the immunoprecipitation had been done with the remainder of the samples: 45 μl of the reactions were mixed with an equal volume of precipitation buffer containing either IgG-agarose beads (I), agarose beads without antibody (B) or protein A sepharose beads with prebound anti-Prp8p antibodies (C) and incubated for precipitation for 2 hours at 4°C. The beads were washed in buffer containing 150 mM NaCl, deproteinised, extracted with phenol/chloroform and the RNAs precipitated. The samples from the immunoprecipitation (coIP) as well as the input samples (splicing) were then resuspended in formamide loading buffer and the precipitated RNAs resolved on a 6% (w/v) polyacrylamide gel. The labelled RNA species of the splicing reactions and of the precipiates were visualised by autoradiography. In two additional samples, recombinant dominant negative Prp2p or Prp16p protein was added to the YMA45/2 extract prior to splicing (+ Prp2^{LAT}p or + Prp16^{AKT}p) and the samples were treated as described above. The positions of the labelled RNA species are indicated: LI-E2, Lariat-intron exon 2; LI, Lariat-intron; E1, exon 1.
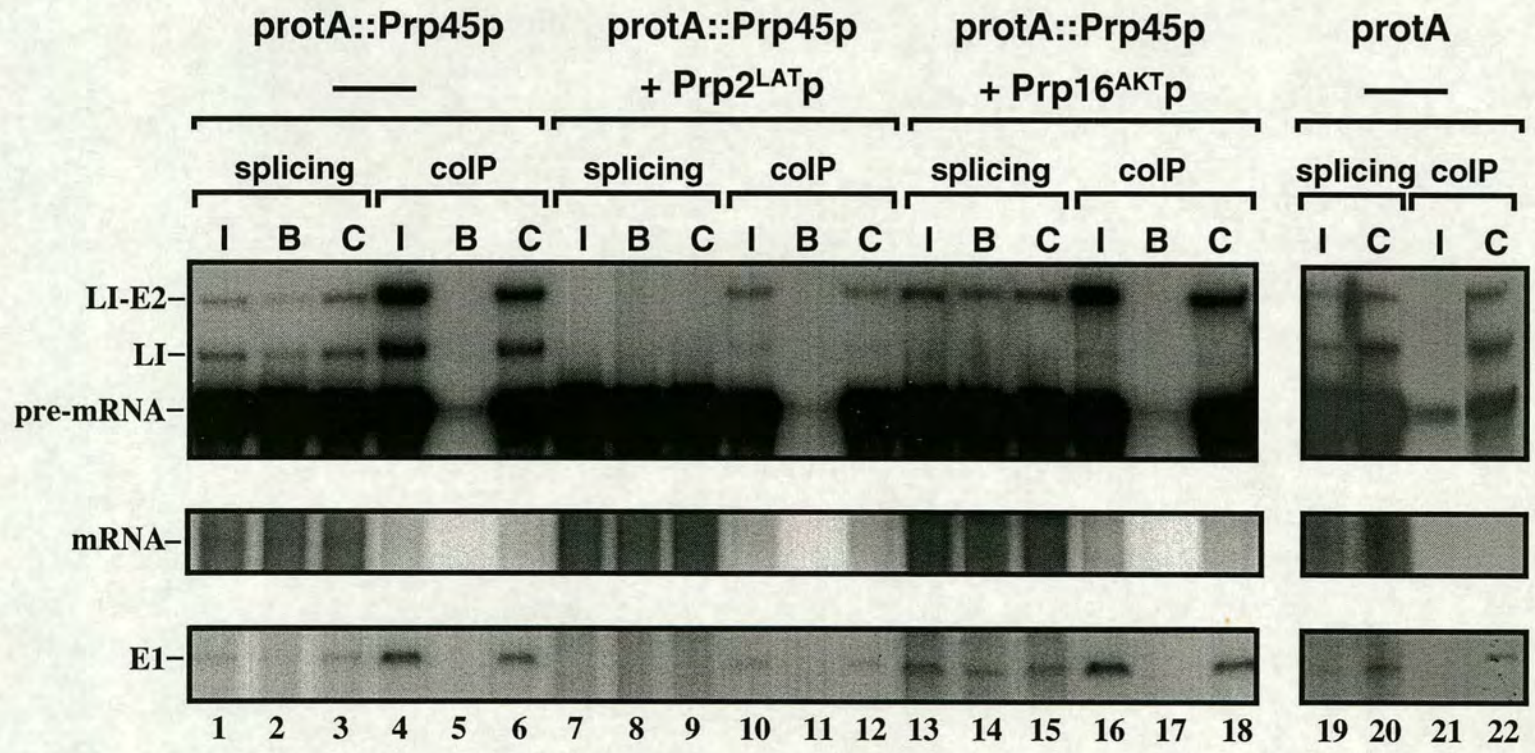
It can be seen that protA::Prp45p efficiently coprecipitated the pre-mRNA, the lariat intron-exon 2 and exon 1 intermediates as well as the excised intron (lane 4). Mature message was not detected in the precipitate. In both negative controls, using agarose without antibody and providing the protein A epitope but no tagged Prp45p, only very small levels of pre-mRNA were found in the precipitate (lane 5 and 21, respectively). These findings showed that Prp45p is associated with the spliceosome throughout the splicing reactions and that it stays associated with the spliceosome until after the second catalytic step, when it is found associated with the excised intron.

When splicing was stalled due to prior addition of Prp2^{LAT}p to the extracts (lanes 7-9), the Prp45 protein still precipitated almost equivalent quantities of pre-mRNA compared to untreated extracts (lane 10 compared to 4). Therefore, Prp45p is already in the spliceosome, before Prp2p joins in to promote the first transesterification reaction. Small amounts of precipitated splicing intermediates, and excised intron from these stalled extracts must be accounted to residual splicing activity of the extracts, due to the presence of wild type Prp2p.

Stalling splicing by addition of dn Prp16^{AKT}p prior to the substrate RNA (lanes 13-18) did not give any more information on the spliceosome association pattern of Prp45p, but it nicely confirmed the data from the co-precipitation experiment of the unstalled reactions: in the extract in which splicing was stalled after the first but prior to the second transesterification, pre-mRNA and the intermediates of the splicing reaction were efficiently precipitated by Prp45p (lane 16).

## VI.6.1 Prp45p coprecipitates spliceosomal snRNAs

To investigate whether Prp45p is associated with any of the spliceosomal snRNAs, again protA::Prp45p was produced in strain YMA45/2 and whole cell extracts (splicing extracts) were prepared. The negative control was strain BMA38n transformed with vector pNOPPATA1L, which produces the double protein A epitope (protA). The extracts were either incubated with immunoglobulin-agarose (IgG-agarose), or as controls with agarose beads or anti-Prp8p antibodies, prebound to protein A-sepharose beads. After the immunoprecipitation, the precipitates were deproteinised, run on an

SDS-polyacrylamide gel and electroblotted. The membrane was subsequently hybridised with radiolabelled oligonucleotides complementary to regions of the U snRNAs. The result of this northern analysis is depicted in figure VI.11.

It was demonstrated that protA::Prp45p co-precipitated significant amounts of U2, U5 and U6 snRNAs but no U1 or U4 snRNAs (lane 6). The agarose beads (no antibody) only precipitated background levels of U1, U5 and U6 snRNAs (lane 7) and the protein A epitope alone also did not precipitate any significant amounts of snRNAs (lane 3). This suggests that Prp45p associates with U2, U5 and U6 snRNAs under non-splicing conditions (no ATP was added to the extracts prior to the precipitation). Taking into account that only 40 % of the amount of extract that was used for the precipitations was used for the preparation of the total RNA, it becomes clear that the association of Prp45p with the U2, U5 and U6 snRNAs is relatively weak. The very efficient precipitation of U4, U5 and U6 snRNAs by antibodies against the tri-snRNP protein Prp8p showed that the precipitation procedure had worked (lanes 5 and 8). Considering that protA::Prp45p could be efficiently precipitated by IgG-agarose, as confirmed by western analysis (data not shown) and furthermore, the fact that it strongly precipitated spliceosomes, suggests that only a small portion of Prp45p is associated with snRNAs under non-splicing conditions.

Figure VI.11:
Coprecipitation of snRNAs by Prp45p. Whole yeast cell extract (splicing extract) was prepared from cells of strain YMA45/2, grown in YMGRSsup liquid medium, producing the protein A tagged Prp45p (protA::Prp45p). As a control, an extract was prepared from strain BMA38n transformed with vector pNOPPATA1L, producing a double protein A epitope (protA). 50 μl of the extracts were mixed with an equal volume of precipitation buffer containing either IgG-agarose beads (I), agarose beads without antibody (B) or protein A sepharose beads with prebound anti-Prp8p antibodies (C) and incubated for precipitation for 2 hours at 4^0C. The beads were washed in buffer containing 150 mM (or 75mM) salt, deproteinised, extracted with phenol/chloroform and the RNAs precipitated. The RNAs were then resuspended in formamide loading buffer, resolved in a denaturing 6% (w/v) polyacrylamide gel and electroblotted to a Hybond-N nylon membrane (Amersham). The RNAs were then probed with radiolabelled oligonucleotides complementary to the spliceosomal snRNAs. The positions of the U1, U2, U4, U5S, U5L and U6 snRNAs are indicated. Total RNA (input): A 20 μl aliquot of the extracts was deproteinised, precipitated and subjected to Northern analysis as the precipitation samples (coIP).
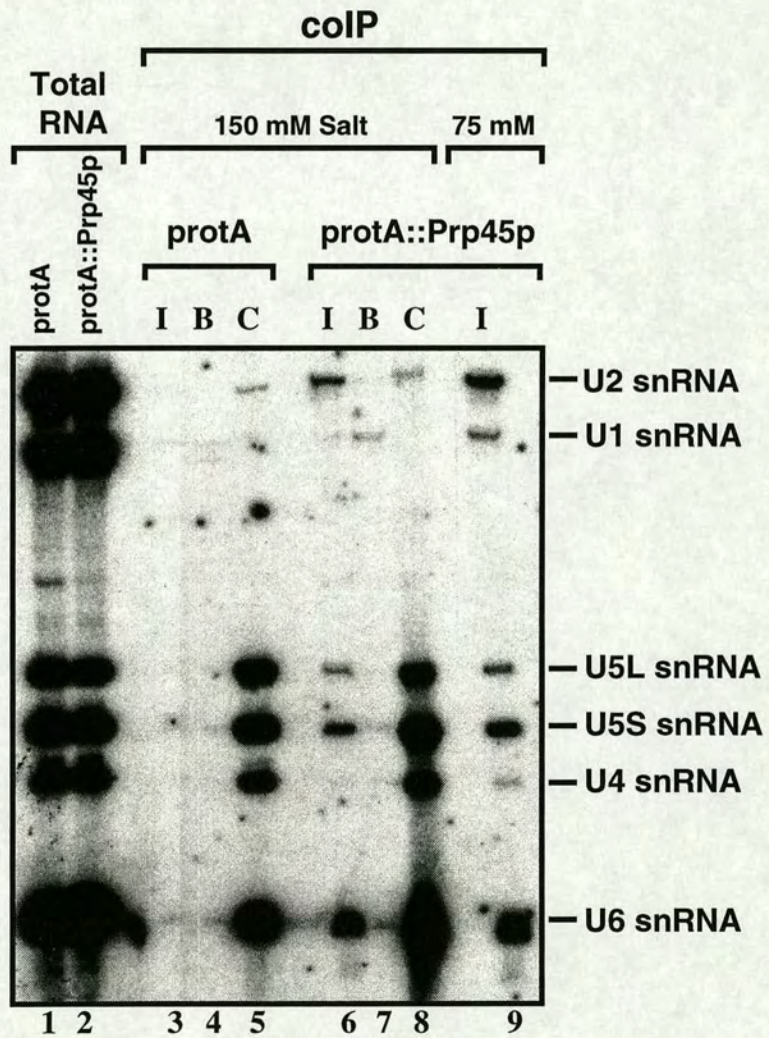
VI.7 Generation of recombinant Prp45p

From the experiments described in the previous sections it was determined that Prp45p enters the spliceosome prior to the first transesterification event, before action of Prp2p, and stays in the spliceosome until both transesterifications reactions are completed. Therefore, it could not be defined, whether Prp45p function was required for the first, the second or both transesterification reactions to proceed. To address this question, it was decided to produce Prp45p in *Escherichia coli* (*E. coli*) in order to use the recombinant protein in depletion/reconstitution experiments of cell extracts produced from strain YMA45/2.

For the production of recombinant Prp45p the pET System (Novagen) was used, in which target genes are cloned under control of the strong bacteriophage T7 promotor. The plasmids are then transformed into expression strains which contain a chromosomal copy of the T7 RNA polymerase gene for the expression of the target genes. Expression of the T7 RNA polymerase gene and consequently expression of the target gene is inducible by addition of isopropyl-β-D-thiogalactoside (IPTG) to the growth medium. The *PRP45* ORF was isolated from the pMA45 vector (section VI.10.1) via *Eco*RI, *Sal*I restriction and recessed 3' termini were subsequently filled in by using the Klenow-fragment of *E. coli* DNA-Polymerase I (Promega). The fragment was then cloned into the *E. coli* expression vector pET19b (Novagen), which had been previously cut with *Nde*I and blunted by Klenow treatment. The resulting construct was then checked by restriction digests and DNA sequencing using an oligonucleotide (T7-1 forward) as primer which allowed sequencing over the 5' junction of the inserted fragment. The plasmid, designated pETMA45 was then transformed into the *E. coli* expression strain BL21 (Novagen). In pETMA45, *PRP45* is under control of the T7 promotor and the ORF is fused n-terminally to a region encoding 10 consecutive histidine residues, which should allow Ni-NTA affinity purification of the produced protein.

After overnight induction of *PRP45* expression in BL21 by addition of 0.75 mM IPTG, the *E. coli* cells were harvested and cell extract was prepared. The extract was then passed over a Ni-NTA column (Qiagen). The column was washed twice with buffer containing first 20 then 55 mM imidazole to wash off any proteins aspecifically

bound to the column. Eventually, proteins still bound to the column were eluted by addition of buffer containing 200 mM imidazole. Aliquots of the fractions (and of the supernate and the washes) were subsequently resolved in a 12% SDS-polyacrylamide gel. Each fraction was run in duplicate on the same gel, which was cut in half after the run. One half of the gel was stained with coomassie blue for detection of the proteins present, the other half was subjected to western blotting to detect specifically the His-tagged Prp45 protein, using mouse monoclonal anti-HIS antibodies (Quiagen). Figure VI.12 shows the coomassie stained gel as well as the result of the western blot.

A strong band in fractions 1 and 2 running at a molecular weight of 50-55 kDa is the main elution product present in the coomassie stained gel. This band also gives a strong signal upon detection by the anti-His antibodies (Quiagen) in the western blot, strongly suggesting that this band represents the His::Prp45 protein. Lower amounts of the protein are also present in fractions 3 and 4. The calculated molecular weight of the His::Prp45p is about 45 kDa. Its decreased electrophoretic mobility could be accounted to its highly charged nature. The finding that the protein runs slower than predicted from the calculation of the molecular weight was actually expected, since as shown in chapter VI.4, a protein A-tagged version of Prp45p also exhibits decreased mobility in SDS-polyacrylamide gels. The relatively strong bands at lower molecular weight which can be seen in fractions 1 and 2 are apparently breakdown- or prematurely terminated products of His::Prp45p, since higher exposures show clearly that their amount decrease in proportion to the His::Prp45p in fractions 2, 3, 4 and 5 (data not shown) and furthermore, no products of that molecular weight can be detected in the supernate.

Fractions 1 and 2 were dialysed to remove the highly concentrated imidazole from the buffer. Aliquots of the dialysed fraction 2, containing purified, recombinant His::Prp45p were used in subsequent experiments.
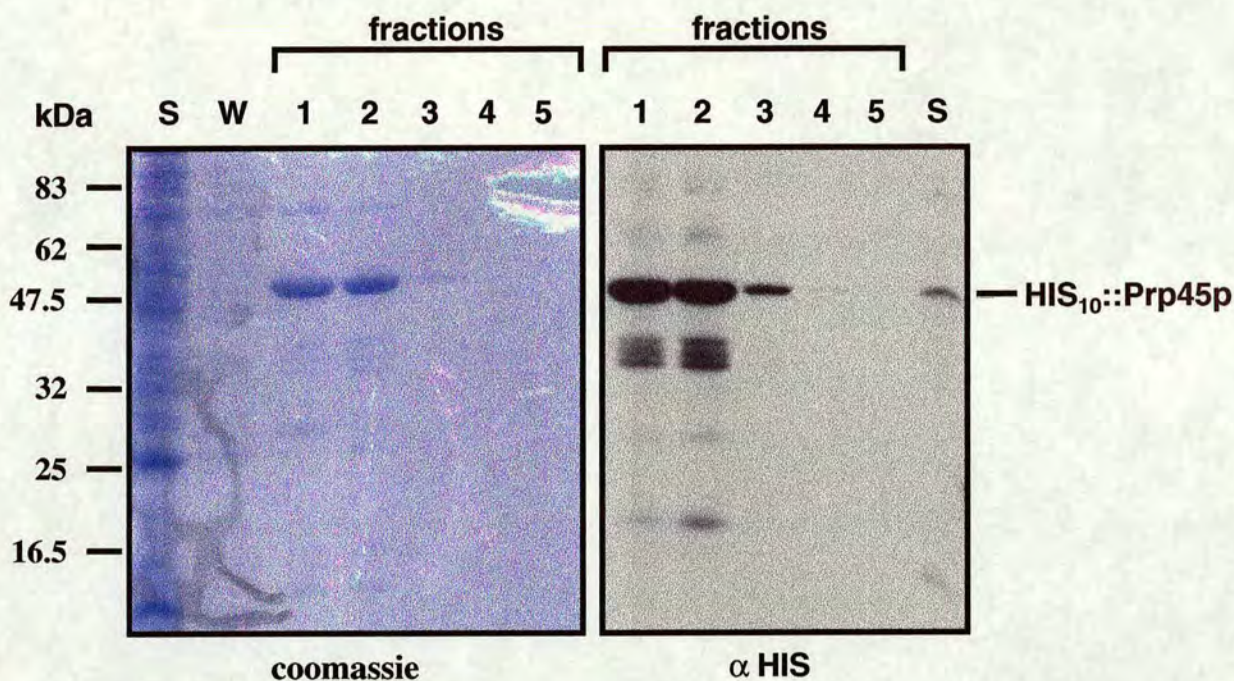
Figure VI.12:
Expression of His-tagged Prp45p in *Escherichia coli*. Two litres of LB
medium containing the appropriate antibiotics were inoculated with *E. coli* strain
BL21, bearing the expression plasmid encoding the His_{10}::Prp45 fusion protein,
pETMA45. The culture was grown at $30^\circ C$ to an OD_{600} of 0.4 and expression
induced by addition of IPTG to a final concentration of 0.75 mM. The culture was
incubated for another 12 hours at $30^\circ C$ and a protein extract prepared. The extract was
incubated with Ni-NTA agarose beads (Qiagen) for 2 hours at $4^\circ C$. The beads were
transferred to a column and washed twice with 10 ml of buffer containing 20 and then
55 mM imidazole. Finally, proteins were eluted in 1.5 ml fractions with buffer
containing 200 mM imidazole. Thirty microlitres of the fractions as well as of the
washes (W, combined) and of the supernatant (S) were then loaded in duplicate and
resolved on a 12% (w/v) SDS polyacrylamide gel. The gel was cut in half with one
part stained with coomassie blue and the other analysed by western blot analysis.
His_{10}::Prp45p was detected with mouse monoclonal anti-penta His antibodies
(Qiagen) and visualised by ECL.

VI.8 Prp45p is required for splicing *in vitro*

As shown in section VI.5, depletion of Prp45p from the cells leads to an accumulation of precursor mRNA, showing that the splicing process is blocked at one point or the other. In order to define whether Prp45p function is required prior to or after the first transesterification reaction, whole cell extracts were prepared from strain YMA45/2 grown under permissive or non-permissive conditions. Then splicing of uniformly radiolabelled actin precursor-mRNA was performed for 30 min at 23°C. Alternatively, 1 μl of fraction 2 (see section VI.7) containing purified His::Prp45p (or 1 μl of a 1:4 dilution of fraction 2) was added to the previously depleted YMA45/2 extract before splicing was performed. The extracts were deproteinised and fractionated by SDS polyacrylamide gel electrophoresis (figure VI.13).

It becomes clear, that the extract from cells depleted of Prp45p is incapable of splicing the actin precursor mRNA (lane 3), whereas the extract from cells grown under permissive conditions (producing protA::Prp45p) splices the precursor to at least 60 % wt levels (lane 1 and data not shown). This finding suggested that Prp45p is required for pre-mRNA splicing *in vitro* and furthermore, since no intermediates are produced, that its function is already required prior to the first transesterification reaction. However, it had to be proved whether the incapability of the Prp45p-depleted extract to splice, was indeed a consequence of Prp45p depletion, or whether it was simply caused by an inadequate preparation of that particular extract. The addition of His::Prp45p indeed restores splicing activity of that extract, although only to some extent (approximately 10-30 % compare lanes 4 and 5 to lane 1). However, the result proves that suboptimal extract preparation is not the cause for the incapability of the extract to splice, but that in fact, the absence of Prp45p prevents splicing in this extract. Why splicing is only partially restored upon addition of recombinant His::Prp45p to the extract could have several reasons. First, maybe in the yeast cell, Prp45p is subjected to posttranslational modifications (such as phosphorylation or glycosylation), which are essential for full activity and which are not found on the recombinant protein. Second, maybe only a small amount of the purified recombinant protein is active. It was noticed that during the dialysis procedure of the isolated

His::Prp45p, the protein precipitated to some extent. Therefore, it could be imagined that a large part of the added protein was actually insoluble and inactive. It seems, though, that the amount of Prp45p is not limiting here, since splicing activity of the extract does not improve by adding four times more Prp45p (compare lanes 4 and 5). The addition of 3 µl undiluted fraction 2 also did not provide an increased splicing activity to the extract (data not shown). Another possible explanation for the poor splicing activity of the reconstituted extract might be that the breakdown products obviously present in fraction 2 (VI.12) interfere with Prp45p function. They may occupy potential binding sites for the full length protein in the spliceosome, but are unable to perform its function, thereby decreasing splicing efficiency. Third, the His-tag could effect Prp45p function.

It was also tried to complement the splicing defect of strain YMA45/2 grown on glucose-based medium by providing full-length Prp45p fused to the LexA protein (LA::Prp45p). It was already shown (chapter VI.4, figure. VI.7) that this does complement the growth defect of the strain under non-permissive conditions. Indeed, also the *in vitro* splicing activity could be restored by producing LA::Prp45p from plasmid pMA45 (lane 2 in figure VI.13). However, again as with the recombinant purified protein, splicing activity can only be partially restored by providing Prp45p. The fact that LA::Prp45p can rescue cell growth to almost wt levels (figure VI.7) but splicing activity *in vitro* only very weakly, suggests that LA::Prp45p is almost fully functional but that the *in vitro* requirements to achieve proper splicing are much more difficult to fulfil. The important qualitative message, however, gained from these depletion/ reconstitution and complementation studies remains: Prp45p is essential for pre-mRNA splicing *in vitro* and its function is required prior to the first transesterification reaction. These results certainly do not exclude the possibility that Prp45p may have additional roles at other stages of the splicing reaction, e.g. during the second step or in spliceosome disassembly as maybe suggested by the interaction with Prp22p.
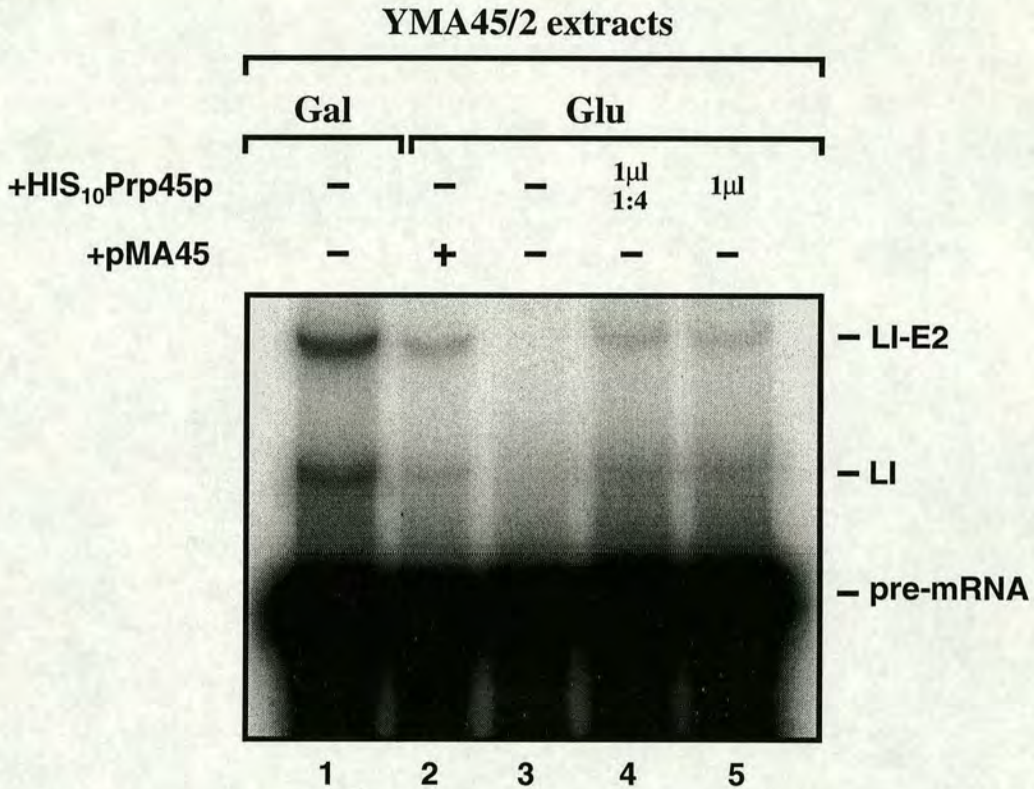
Figure VI. 13:
Complementation of Prp45p-depleted splicing extracts. Whole cell extracts (splicing extracts) were prepared from cells of strain YMA45/2 grown for 12 hours to an OD_{600} of 1.3 in either YMGRSsup (Gal) or YPDA (Glu) liquid medium. Splicing of radiolabelled actin pre-mRNA was performed with 10 μl (total volume 20 μl) of neat extracts as well as with the extracts after prior addition of His_{10}::Prp45p recombinant protein. Splicing was also performed with an extract from YMA45/2 cells grown in YMM -W, which had been transformed with plasmid pMA45, coding for a LexA::Prp45 fusion protein. After splicing, the reactions were deproteinised, extracted with phenol/chloroform, the RNAs precipitated and resolved in a denaturing 6% polyacrylamide gel. The result was visualised by autoradiography. The positions of the RNA species are indicated: LI-E2, Lariat-intron exon 2; LI, Lariat-intron. Gal - YMGRSsup; Glu - YPDA or YMM-W; 1 μl - one microlitre of fraction 2 (figure VI. 13) after dialysis was added; 1:4 - 1 μl of fraction 2 after dialysis was diluted 1:4 in splicing buffer and added prior to splicing.

VI.9 Functional homology between Prp45p and the human SNW1 protein

In section VI.2 the primary structure similarity of Prp45p to a number of proteins in other organisms was presented. The human protein SNW1 (also called Skip or NCoA-62) which shares 22.8% sequence identity with Prp45p was demonstrated to be a component of the spliceosome (Neubauer *et al.*, 1998). In order to test, whether Prp45p and SNW1 are in fact functional homologues, complementation of the growth defect of strain YMA45/2 (depleted of Prp45p) was attempted by overexpression of *SNW1*.

Therefore, strain YMA45/2 was transformed with plasmid pGBT9/SNW1. This two-hybrid bait plasmid produces the DNA-binding domain of Gal4p fused to the SNW1 protein (GBD::SNW1). Plasmid pGBT9/SNW1 was kindly provided by Mike Hayman. As a control, plasmid pGBT9 (producing GBDp) was transformed in parallel. Figure VI.14 shows growth of strain YMA45/2 containing pGBT9 or pGBT9/SNW1 under permissive and repressing conditions.

It is clearly demonstrated here that overexpression of the GBD::SNW1 protein in YMA45/2 cells overcomes the growth defect on glucose-based medium. When the GBDp alone is produced, only a background level of growth can be detected. Although it seems that the human SNW1 protein does not complement the growth defect as efficiently as the yeast Prp45p (LA::Prp45p), it can definitely be concluded that SNW1 can to some extent substitute for the function of Prp45p in the yeast cells. This means that the human protein SNW1 (Skip, NCoA-62) represents the functional homologue of Prp45p, and therefore will be henceforth referred to as human Prp45p (hPrp45p).

The fact, that hPrp45p and Prp45p are functional homologues, although the amino acid identity between the proteins is only 22.8%, strongly suggests that also the SNW proteins of *S. pombe*, *A. thaliana*, *D. melanogaster*, *P. falciparum* and *D. discoideum*, which share a much higher degree of homology to hPrp45p are pre-mRNA splicing factors. It remains to be determined whether maybe additional proteins in yeast compensate for the probable evolutionary loss of some domains in Prp45p, which are highly conserved in the other SNW proteins. Alternatively, some domains

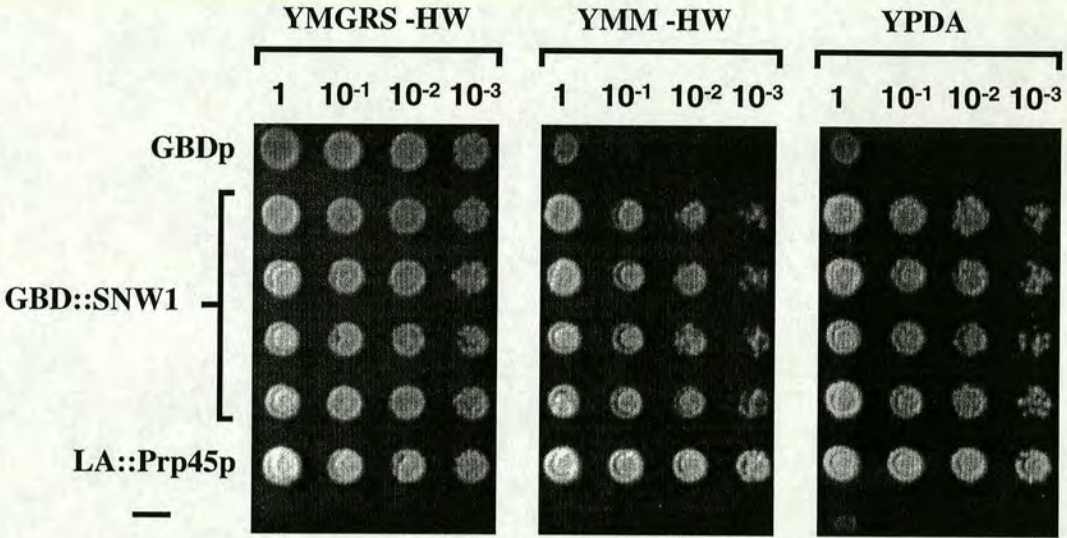might be redundant in yeast because regulatory functions (such as alternative splicing) are not required.

Figure VI.14:
Complementation of Prp45p-depletion by the human SNW1 protein.
Yeast strain YMA45/2 was transformed with either pGBT9 (producing the Gal4p DNA-binding domain, GBDp), pGBT9/SNW1 (producing the human protein SNW1, n-terminally fused to GBDp, GBD::SNW1) or pMA45 (producing Prp45p fused to the LexA protein, LA::Prp45p). Transformants were selected on YMGRS -HW medium, suspended in serial dilutions in microtiterplates and then transferred onto either galactose- (YMGRS -HW) or glucose- (YMM -HW, YPDA) based media. Four independent transformants producing the GBD::SNW1 protein are depicted. - the bottom lane shows the untransformed YMA45/2 strain. The plates were incubated for 2.5 days at 30°C.

VI.10 The Prp45p two-hybrid screen

In the previous sections it was demonstrated that Prp45p is a novel essential splicing factor. This section will describe the performance of a two-hybrid screen with Prp45p as bait. It was hoped to gain results which would confirm the Prp22p-Prp45p

interaction and maybe reveal additional interactions of the novel splicing protein with the splicing machinery.

VI.10.1 Cloning the Prp45p bait vector

In order to construct a bait vector containing the complete *PRP45* ORF, oligonucleotide primers Func1 and Func2 were used to amplify the *PRP45* ORF from genomic DNA of strain BMA38n. The PCR product was gel purified using the Quiagen gel extraction kit and subsequently cut with the restriction enzymes *Eco*RI and *Sal*I. The resulting product was cloned into vector pBTM116, which had been cut with *Eco*RI and *Sal*I Then the product was sequenced using oligonucleotides W2248 (LexA forward) and LexA reverse. The complete ORF was sequenced twice, and no deviations from the sequence in the database were found. The vector, designated as pMA45, was transformed into yeast strain L40. pMA45 should express Prp45p fused to the LexA protein (LA::Prp45p). The expression and functionality of the bait fusion protein was tested via transformation of pMA45 into strain YMA45/2, which contains a conditionally regulated *PRP45* allele. The plasmid was able to complement the growth defect of YMA45/2 under non-permissive conditions (chapter VI.4, fig. VI.7) and furthermore it was able to partially complement the splicing deficiency of a splicing extract prepared from YMA45/2 cells grown under repressing conditions (section VI.8, figure VI.12).

VI.10.2 The Prp45p two-hybrid screen

Prp45p bait culture (plasmid pMA45 in yeast strain L40) was grown overnight to an OD_{600} of 1.1 and then mixed with 1 ml of FRYL in 20 ml of YPDA. The haploids were mated for 4 hours on filters and subsequently collected and spread onto 105 YMM -LWH plates. The haploid strains mated with an efficiency of 17 % and a total of 34 million diploids were screened for expression of the *HIS3* reporter gene. After 3 days approximately 2000 colonies had grown on the selective plates. 1542 large and middle

sized colonies were picked, suspended in microtiter plates and then spotted onto YMM -LWH plates containing either 5 or 20 mM 3AT to increase the stringency on the reporter gene expression. After 3 days, 115 colonies exhibited significant growth on 20 mM 3AT, another 192 colonies grew on plates containing 5 mM 3AT. It was therefore decided to analyse the 115 clones, which obviously contained the strongest interacting preys. After purification of the colonies by restreaking them 3 times on YMM -LWH medium, the prey inserts were amplified by yeast colony-PCR. The resulting PCR products were gel-purified using the QIAquick Gel Extraction Kit (Qiagen) and their identity was determined by DNA-sequencing using oligonucleotide W2248 (lexA forward). In some cases lexA reverse primer was used in addition, to determine the exact 3' end of the inserts. Table VI.2 summarises the results of the Prp45p two-hybrid screen.

Nine prey proteins that can be categorised as A1 candidates were found to interact with Prp45p:

ORF *YDR141c* encodes a 194 kDa hypothetical protein, which was previously classified as a putative integral membrane protein, due to the presence of several strongly hydrophobic domains in its primary structure (Nelissen *et al.*, 1997). The protein was found only twice in the screen.

The Fir1 protein, encoded by ORF *YER032c*, was found in two different fusions a total of four times. Fir1p function obviously influences polyadenylation-cleavage efficiency. The protein has already been discussed in some detail in chapter IV.3, because it was also found in a two-hybrid screen, using Ymr44p as bait, which was identified in turn as a Prp18p two-hybrid interactor. The two small c-terminal fragments of Fir1p, which were found to interact with Prp45p, had also been identified with the Ymr44p bait. This could mean that either Prp45p and Ymr44p bind successively to this c-terminal fragment or alternatively, Prp45p, Fir1p and Ymr44p act in a complex, which might connect the splicing and polyadenylation processes. However, as discussed in chapter IV.3, the connection of Ymr44p to the splicing pathway is somewhat dubious, since the protein interacts with splicing factors associated with U1, U2 and U5 snRNPs and a sensible explanation of Ymr44p

function to link splicing and polyadenylation is therefore difficult to propose. In addition, the Fir1p was shown to interact aspecifically with a number of baits unrelated to splicing and polyadenylation in directed two-hybrid assays (del Olmo *et al.*, 1997) and therefore it is questionable whether the Prp45p-Fir1p interaction is indeed of biological relevance or whether the c-terminus of Fir1p is simply predisposed to cause false positive two-hybrid results. Additional investigations need to be done, preferably genetic assays, which do not depend directly on sequence and structure of the participating proteins, which could demonstrate connected functions of these splicing and polyadenylation factors.

Three different fragments of a hypothetical 95 kDa protein encoded by ORF *YGR198w* were found to interact with the Prp45p bait. No function has as yet been assigned to the protein and it does not share significant homology to other proteins in the databases.

The ORF *YGR270w* encodes the 157 kDa non-essential protein Yta7p. Yta7p was found in two independent fusions a total of four times in the screen. The protein belongs to a family of ATPases, whose members play diverse roles in cell cycle control, promotion of organelle fusion, protein degradation and gene expression (e.g. Tomoyasu *et al.*, 1993; Thorsness *et al.*, 1993). The cellular role of Yta7p is not known.

The ORF *YLR024c* encodes a hypothetical 216 kDa protein, which shows similarity to the yeast ubiquitin-protein ligase (E3) Ubr1p. Its function is not known to date. The protein was found 4 times in two independent fusions in the screen.

Table VI.2:
Results of the Prp45p two-hybrid screen. A LexA-Prp45p fusion protein (LA::Prp45p) was used as bait to screen the FRYL for interacting proteins. Preys of the A categories are ordered alphabetically according to their ORF name as defined in the *Saccharomyces* Genome Database, SGD. B category preys are listed at the end of the table. No. - frequency, with which a fragment was identified in the screen; Chr - chromosome number; Strand: w - Watson DNA strand, c - Crick DNA strand (as defined by SGD); nt. from AUG - number of nucleotide at which fusion starts (A from the initiation codon AUG is 1); Insert size - insert length (nucleotides) determined by sequencing, * - insert size has been determined only approximately via *Bam*HI digest of the prey plasmid and subsequent agarose electrophoresis. Preys were classified in categories A1, A2, A3, A4 and B as defined by Fromont-Racine *et al.*, 1997 and described in section III.1.

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size (bp)	Insert size (bp)	Category	Protein info
28	2	CDC27	YBL084c	II	c	90	2277	1148	A2	Component of anaphase promot. complex, 8 TPR-motifs
27	1	PKC1	YBL105c	II	c	901	3456	1000*	A3	Put. protein kinase
12	1	RAD16	YBR114w	II	w	175	2373	400*	A4	DNA-helicase, radiation repair
106	1		YDL001w	IV	w	1084	1293	1200*	A4	
68	1	NUP84	YDL116w	IV	w	508	2181	600*	A4	Homology to mammalian Nup107p
36n	1		YDR141c	IV	c	1738	5097	n.d	A1	
36c	1		YDR489w	IV	w	n.d	885	n.d		
50	1		YDR141c	IV	c	1399	5097	1782	A1	
43	1	NGG1	YDR176w	IV	w	1524	2109	400*	A4	Transcription factor, histone acetyltransferase complex
51	2	SYF1	YDR416w	IV	w	1325	2580	1157	A3	Splicing factor, syn. lethal with cdc40
76	1	FIR1	YER032w	V	w	2125	2778	440	A1	3' end processing, polyA-cleavage site efficiency
13	3	FIR1	YER032w	V	w	1957	2778	539	A1	3' end processing, polyA-cleavage site efficiency
39	1	TRP5	YGL026c	VII	c	1850	2124	400*	A4	Tryptophan synthetase
65	1	ADE6	YGR061c	VII	c	2374	4097	1200*	A3	5'-phosphoribosylformyl-glycinamidine synthetase
61	1		YGR071c	VII	c	903	2588	1900*	A3	
72	1		YGR198w	VII	w	1351	2454	1071	A1	
78	1		YGR198w	VII	w	2233	2454	527	A1	
90	1		YGR198w	VII	w	2242	2454	534	A1	
85	2	YTA7	YGR270w	VII	w	2437	4140	1156	A1	Member of cdc48/PAS1/SEC18 family of ATPases
74n	2	YTA7	YGR270w	VII	w	2503	4140	1300	A1	Member of cdc48/PAS1/SEC18 family of ATPases
54	1	YAP1801	YHR161c	VIII	w	334	1914	300*	A4	Yeast assembly protein, binds Pan1p and clathrin
57	1	SEC6	YIL068c	IX	c	298	2418	1500*	A3	Component of the exocyst component
58	1		YJL103c	X	c	655	1857	400*	A4	
8	1		YJL149w	X	w	1522	1993	400*	A4	
109	1	APL1	YJR005w	X	w	475	2103	1100*	A3	Beta-adaptin, clathrin-associated complex
100	1		YJR061w	X	w	1183	2809	1500*	A3	
71	2	DYN1	YKR054c	XI	c	2068	12279	2200*	A3	Heavy chain of cytoplasmic dynein
11n	1		YLR024c	XII	c	2458	5620	900*	A1	Probable membrane prot., ubiquitin-ligase (E3) like
11c	1	CDC27	YBL084c	II	c				A	
47n	3		YLR024c	XII	c	2500	5620	1800*	A1	Probable membrane prot., ubiquitin-ligase (E3) like
47c	1	(SSP2)	(YOR242c)	XV	w				B	
4	3	SYF3	YLR117c	XII	c	101	2064	860	A1	Splicing factor, hom. to D.melanogaster crn-protein
22	3	SYF3	YLR117c	XII	c	32	2064	436	A1	Splicing factor, hom. to D.melanogaster crn-protein
69	3		YLR187w	XII	w	498	3081	400*	A4	
84	2		YLR320w	XII	w	1831	4365	1200*	A3	
63	3		YLR386w	XII	w	798	2643	1000*	A1	
113	2		YLR386w	XII	w	991	2643	800*	A1	
9	1		YLR424w	XII	w	93	2187	1200*	A2/3	
111	1	NUP188	YML103c	XIII	c	3871	4968	1100*	A3	Nucleoporin
67	1		YMR031c	XIII	c	-208	2532	437+	A2	
82	1	ECM5	YMR176w	XIII	w	2368	4236	1600*	A3	Extra cellular mutant
66	1	SGS1	YMR190c	XIII	c	1507	4344	709	A1	DNA helicase motifs, genome stability
88	1	SGS1	YMR190c	XIII	c	1198	4344	1088	A1	DNA helicase motifs, genome stability
92	1		YNL078w	XIV	w	809	1224	1500*	A4	

...continued

Clone	No.	Gene	ORF	Chr	Strand	nt. from AUG	ORF size (bp)	Insert size (bp)	Category	Protein info
49	1	SLA2	YNL243w	XIV	w	340	2907	1700*	A3	Transmembrane protein, cell polarization and endocytosis
70	1		YNL246w	XIV	w	77	888	440+	A2	
30	1	BNI1	YNL271c	XIV	c	4441	5862	700*	A4	Formin homology domain, homol. to BNR1
64	1		YNR059w	XIV	w	583	1743	400*	A4	
89	1		YOR023c	XV	c	55	1701	850*	A2	
62	1	NIF1	YOR156c	XV	c	1082	2181	900*	A4	Interacts with cdc12 (in two-hybrid)
59	2		YPL151c	XVI	c	379	1356	1076	A1	Splicing factor (this work)
56	1		YPL151c	XVI	c	-66	1356	1802	A1	Splicing factor (this work)
1	2		YPL151c	XVI	c	51	1356	1495	A1	Splicing factor (this work)
3	11		YPL151c	XVI	c	69	1356	1229	A1	Splicing factor (this work)
31	2		YPL151c	XVI	c	45	1356	1894	A1	Splicing factor (this work)
45	1		YPL151c	XVI	c	96	1356	1309	A1	Splicing factor (this work)
105	1		YPR023c	XVI	c	682	1206	500*	A4	
16	1	SKI3	YPR189w	XVI	w	1387	4299	1200*	A3	Antiviral prot., contains 8 TPR-motifs
6	1	anti	(YAL017w)	I	c			1200*	B	
99	1	anti	(YAL024c)	I	w			1100*	B	
93	1	anti	(YBR156c)	II	w			750*	B	
18	1	anti	(YGL064c)	VII	w			1000*	B	
19	1	anti	(YGR059w)	VII	c			1100*	B	
80	1	anti	(YIL068c)	IX	w			750*	B	
44	2	anti	(YJR109c)	X	w			2400*	B	
25	2	anti	(YJR110w)	X	c			3000*	B	
53	1	anti	(YLR187w)	XII	c			300*	B	
74c	1	anti	(YLR439w)	XII	c			n.d	B	
21	3	anti	(YML083c)	XIII	w			800*	B	
103	1	anti	(YNL140c)	XIV	w			1300*	B	
96	1	anti	(YOL022c)	XV	w			n.d	B	
48	2	anti	(YPL022w)	XVI	c			200*	B	
34	1	anti	(YPL152w)	XVI	c			500*	B	
40	1	anti	(YPL152w)	XVI	c			1000*	B	
29	2		YHRC Ty1-1					n.d	B	
23	1	mito						800*	B	

The Syf3 protein, a recently described splicing factor, also called Clf1p (Chung *et al.*, 1999), encoded by ORF *YLR117c*, was found six times in two different fusions (figure VI.15).
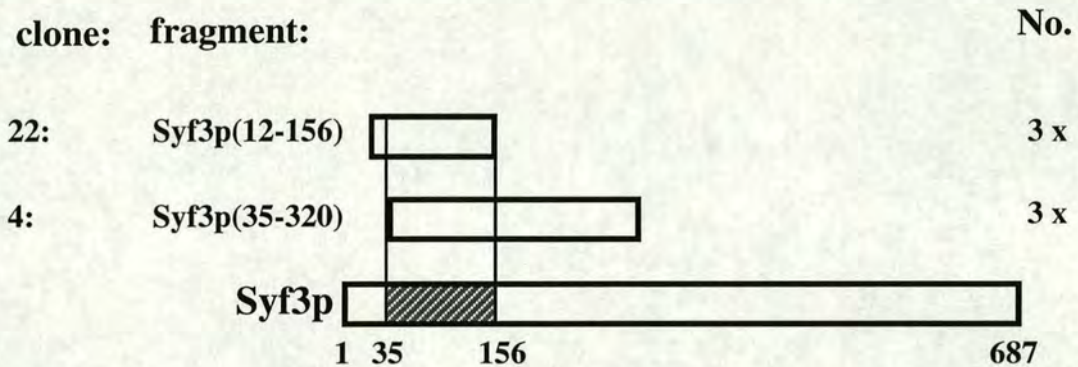

Figure VI.15:
Syf3p prey fragments identified in the Prp45p two-hybrid screen. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen.

This result reciprocates the interaction seen when Syf3p was used as bait in a two-hybrid screen: Prp45p was identified as an A2 prey candidate with the Syf3 bait (Ian Dix and Caroline Russell, unpublished data). Intriguingly, Syf3p was also identified as interacting factor of the A1 category in the Prp22p two-hybrid screen (chapter V.3). The Syf3p prey fragments found to interact with Prp45p cover a region close to the n-terminus of the protein, whereas the fragments interacting with Prp22p share a common region at the c-terminus, suggesting that the recognised two-hybrid interactions are not due to a general "stickiness" of a particular region within the Syf3 protein. The characteristics and known functions of the Syf3 protein, as well as the significance of the strong two-hybrid interactions of Prp22p with Prp45p and Syf3p, and of Prp45p with Syf3p, will be the subject of discussion at the end of this chapter.

The 100 kDa protein Ylr386p was identified in two prey fusions a total of 5 times. Ylr386p had previously been identified as significant interactor in a two-hybrid screen with the U6 snRNP core protein Lsm4p and was studied in some detail in this context (Andrew Mayes, PhD thesis, 1999). The protein was shown not to be essential for cell viability, but disruption of the ORF leads to a slow growth phenotype at 37 °C. A link with the splicing machinery could not be confirmed in the investigations done: inactivation of Ylr386p function did not affect the cellular levels of snRNAs (in contrast to Lsm4p, which is required for maintenance of U6 snRNA levels in the cell; Mayes *et al.*, 1999). Nor did a two-hybrid screen with Ylr386p as bait reveal any interactions which would suggest an involvement of Ylr386p in splicing (Andrew Mayes, PhD thesis, 1999). In addition, when one compares the preys isolated with the Ylr386p bait with the preys found in the Prp45p screen, there are no overlaps present at all, making an involvement of Ylr386p and Prp45p in the same cellular process, or even a biological relevant interaction between the proteins unlikely.

Two clones were isolated, which contained prey plasmids encoding independent fusions of the Sgs1 protein (ORF *YMR190c*). The 163 kDa protein is an ATPase of the DEAD-box family, binds DNA and RNA, has a DNA-helicase activity and is thought to act as gyrase in concert with DNA topoisomerase III (Gangloff *et al.*, 1994; Ng *et al.*, 1999 and references therein). The protein was also found as an interacting protein of the A3 category in a two-hybrid screen with the second step splicing factor and cell cycle protein Prp17p (A. Colley, personal communication).

The statistically most significant interacting protein found in the Prp45p screen, however, is encoded by ORF *YPL151c*. The protein was found in six different prey fusions a total of 19 times, suggesting a strong and possibly biologically important interaction with Prp45p (see figure VI.16). Interestingly, the common domain of the Ypl151p prey fragments (amino acids 127-432) consists entirely of 7 consecutive WD repeat units, a peptide motif found in many regulatory proteins (Smith *et al.*, 1999). This suggests that one or more WD repeats mediate the interaction with Prp45p. It is likely that all 7 repeats are required, maybe in order to form a functional structure, since no smaller prey fragments have been found.
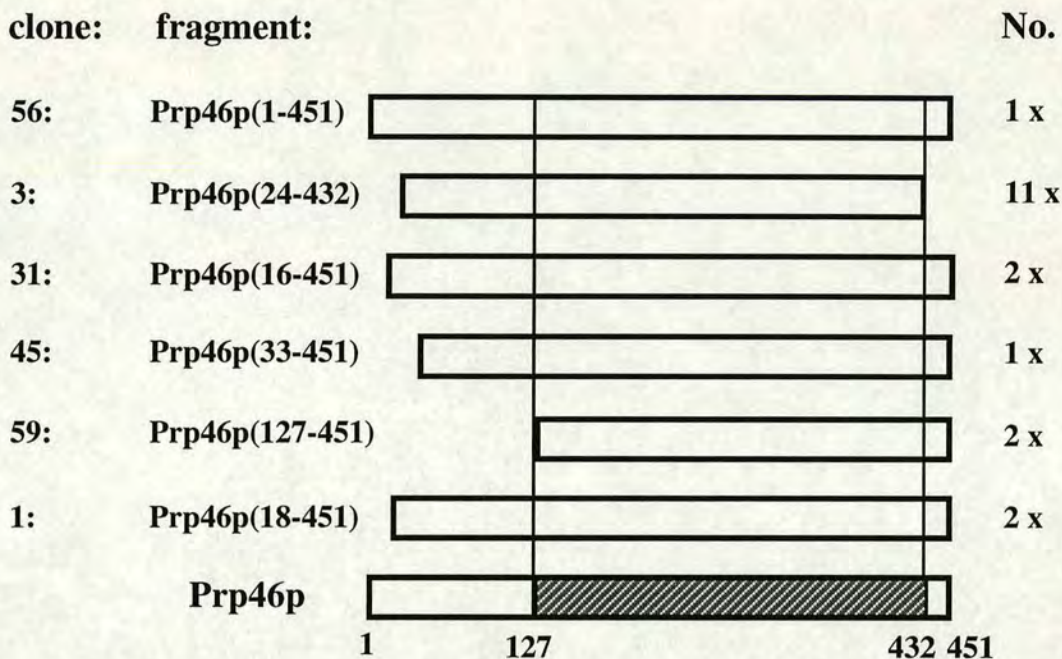
clone:	fragment:		No.
56:	Prp46p(1-451)		1 x
3:	Prp46p(24-432)		11 x
31:	Prp46p(16-451)		2 x
45:	Prp46p(33-451)		1 x
59:	Prp46p(127-451)		2 x
1:	Prp46p(18-451)		2 x
	Prp46p		

1   127   432 451

Figure VI.16:
Prp46p prey fragments identified in the Prp45p two-hybrid screen. The common domain of the prey fragments is indicated as a shaded box in the full length protein. Numbers in brackets describe the amino acid boundaries of the fragment. No. - frequency with which a prey fragment was identified in the screen.

Intriguingly, Ypl151p was also found as an A2 candidate in the two-hybrid screen with the Syf3 protein bait (Ian Dix and Caroline Russell, unpublished data), adding weight to the hypothesis that Ypl151p functions in a complex with Prp45p and Syf3p. Some initial functional characterisation has been done in this work to determine whether Ypl151p is involved in splicing. The data indeed strongly suggest a requirement of the protein for processing of pre-mRNA. Ypl151p was therefore renamed Prp46p, to indicate its involvement in splicing. Chapter VII is dedicated to Prp46p and will contain a detailed presentation of its primary structural features, its homologues and the functional analyses performed.

Another potentially meaningful prey candidate is the 100 kDa Syf1 protein, encoded by ORF*YDR416w*, which was found twice as A3 candidate. Syf1p, as Syf3p, was originally isolated in a screen for proteins being synthetic lethal with a deletion mutant of the second step splicing- and cell cycle factor Prp17/Cdc40 (Ben-Yehuda and Kupiec, unpublished results), suggesting a closely linked function of the proteins in splicing and/or the cell cycle. Indeed, it has been shown in this laboratory, that Syf1p is an essential splicing factor required for splicing before the first transesterification reaction takes place (Caroline Russell, unpublished results). Syf1p and Syf3p are structurally related in that they both contain so called tetratricopeptide (TPR) motifs, which have been implicated in protein-protein interaction (discussed in section VI.10.3). Both proteins have been demonstrated to be essential for the cells to progress from the G2-phase of the cell cycle to mitosis (Ben-Yehuda and Kupiec, unpublished results). Besides the identification of Syf1p and Syf3p in the Prp45p screen, the two proteins have been identified together in two additional two-hybrid screens performed in this laboratory, namely with the Isy1 protein and the Syf2 protein (Ian Dix and Caroline Russell, personal communication; Dix *et al.*, 1999). The above data strongly suggest, that Syf1p and Syf3p act in close proximity to each other and that the observed two-hybrid interactions of these proteins with Prp45p reflect a concerted action of the three proteins at a particular stage of splicing.

In summary, the Prp45p two-hybrid screen led to the identification of a novel splicing protein Prp46p (chapter VII). Furthermore, interactions with two recently characterised splicing factors, Syf1p and Syf3p, were revealed. Unfortunately, the strong two-hybrid interaction of Prp22p bait with Prp45p prey was not directly confirmed in the reciprocal orientation. However, the identification of a two-hybrid interaction between Prp45p and Syf3p indirectly confirms at least a close proximity of Prp45p with Prp22p, since Prp22p bait isolated both Prp45p and Syf3p (chapter V). Furthermore, a weak but significant two-hybrid interaction was observed, when the Prp22p-interacting region of Prp45p was produced from the bait vector and tested against full length Prp22p prey in a directed two-hybrid assay (see chapter VI.3, figure VI.3B). This suggests that Prp45p interacts with Prp22p, but that maybe a large fragment of Prp22p is needed for the structural formation of the interaction domain. A

sufficiently large fragment may not be present in the two-hybrid library which has an average insert size of 700 bp.

VI.10.3 The TPR motif protein Syf3p might form a scaffold for a multi protein complex

In the Prp45p- as well as in the Prp22p two-hybrid screen the Syf3 protein was identified as A1 prey candidate. The finding that Prp22p also identified Prp45p as A1 candidate, supports the hypothesis that these proteins act in a complex. Indeed Syf3p belongs to a family of proteins which contain a degenerate 34 amino acid motif that promotes protein-protein interactions: the tetratricopeptide (TPR) motif (for a review see Lamb *et al.*, 1995). The TPR motif is defined by 8 loosely conserved amino acid residues, namely -W-LG-Y-A-F-A-P, which are found at conserved positions within the motif. The motif is normally present as tandem arrays and is repeated between 3 and 16 times within a protein. Das *et al.* (1998) solved the crystal structure of the three TPR motif-containing protein phosphatase 5 and were able to demonstrate that adjacent TPR motifs are packed together in a parallel arrangement such that a tandem TPR motif structure is composed of a regular series of antiparallel α-helices. They suggest that this arrangement of neighbouring α-helices defines a helical structure and an amphipatic groove and they predict, that multiple-motif TPR proteins would fold into a right handed super-helical structure with a continuous helical groove suitable for the recognition of target proteins. Indeed, a significant number of TPR-containing proteins can be found in multiprotein complexes in which the TPR motifs are thought to form scaffolds to mediate protein-protein interactions. Examples of such complexes are the anaphase promoting complex, containing the TPR motif proteins Cdc16, Cdc23 and Cdc27, the mitochondrial import receptor complex (MAS70), the peroxisomal import receptor complex (PAS8, PAS10 and PXR1) and the transcription repression complex which contains the Cyc8 TPR motif protein (for references see Goebl and Yanagida, 1991 and Lamb *et al.*, 1995 and references therein).

228

Syf3p consists of 15 consecutive TPR repeats and is therefore a good candidate to mediate protein-protein interactions within a larger complex. Indeed, the multiple two-hybrid interactions involving Syf3p suggest that the protein could serve as a scaffold for the formation of a large protein complex. Syf1p, containing 9 TPR repeat units, could play a similar, or supporting role in this putative complex. Figure VI.17 shows the modular organisation of Syf3p and furthermore indicates the putative sites of interaction not only with Prp22p and Prp45p, but also with five other proteins identified as interacting proteins in two-hybrid screens.

Prp45p interacts with an n-terminal region of Syf3p, which covers most of TPR motif 1, motifs 2 and 3 and approximately half of motif 4. Three additional proteins interact with the Syf3 protein in essentially this region: Cef1p, Isy1p and Prp46p. The Cef1 protein is an essential splicing factor, which was isolated as a component of a protein complex associated with the splicing factor Prp19p (Tsai *et al.*, 1999). Like Prp19p, Cef1p joins the spliceosome after the addition of the U2 snRNP and concomitantly with or immediately after dissociation of the U4 snRNP, i.e. just before the formation of an active spliceosome. Cef1p was also isolated in two-hybrid screens performed in this laboratory with the Isy1p bait, the Syf2p bait and the Syf3p bait (all interactions were A1 category; Ian Dix and Caroline Russell, unpublished results). The Isy1 protein was found as prey in two-hybrid screens with the Syf1p bait (A1) and with the Syf3p bait (A4; Dix and Russell, unpublished). Isy1p is non essential for cell viability but its deletion leads to a mild splicing defect *in vivo* (Dix *et al.*, 1999). The authors showed that the protein is associated with the spliceosome throughout the splicing reactions and that it is weakly associated with U5 and U6 snRNAs. Prp46p, originally identified with the Prp45p bait (see chapter VI.10), also interacts with the Syf3p when used as bait (A1 interaction; this work, M. Mellor-Clark, A. Clark and C. Russell). This two-hybrid screen furthermore reciprocated the interaction of Prp45p with Prp46p, since Prp46p bait isolated three independent fusions of Prp45p.

Three proteins putatively interact with regions further to the c-terminus of Syf3p: the Syf2 protein interacts within the region of TPR motifs 10-14; the Isy1 protein seems to have a second site of interaction within Syf3p, located somewhere in the region of TPR motifs 7-14; and a protein named Ntc20p interacted with prey fragments that

overlap in the region of TPR motifs 9-14 (Dix and Russell, unpublished; Mellor-Clark and Russell, personal communication).

The *SYF2* gene, like *SYF1* and *SYF3*, was originally identified in the already mentioned screen for mutants which exhibited synthetic lethality with a deletion of the gene encoding the cell cycle and splicing protein Cdc40p/Prp17p (Ben-Yehuda and Kupiec, unpublished results). In this laboratory it was shown that Syf2p is associated with the spliceosome and that it weakly coprecipitates the U5 and U6 snRNAs (Russell, unpublished data). However, unlike Syf1p and Syf3p, the protein is not essential for cell viability, nor does a complete deletion of the gene lead to any recognisable growth- or splicing defects. When used as bait in a two-hybrid screen, it not only interacted significantly with Syf3p, but also with Cef1p and Syf1p (all A1 interactions; Dix and Russell, unpublished data).

The Ntc20 protein (Ntc20: nineteen complex, 20 kDa) was also shown to be part of the above mentioned Prp19p-associated complex (Tsai *et al.*, 1999). Ntc20p associates with the spliceosome in the same manner, i.e. at the same timepoint as Cef1p and Prp19p, which led to the suggestion that the whole Prp19p complex is added to the spliceosome as an integral, preformed complex. In addition to Syf3p, the Ntc20p two-hybrid screen identified Syf1p and Cef1p as A1 and A3 interacting proteins, respectively (Mellor-Clark and Russell, personal communication). The Syf3p bait identified Ntc20p as interacting prey of the A1 category, thereby reciprocating the interaction.
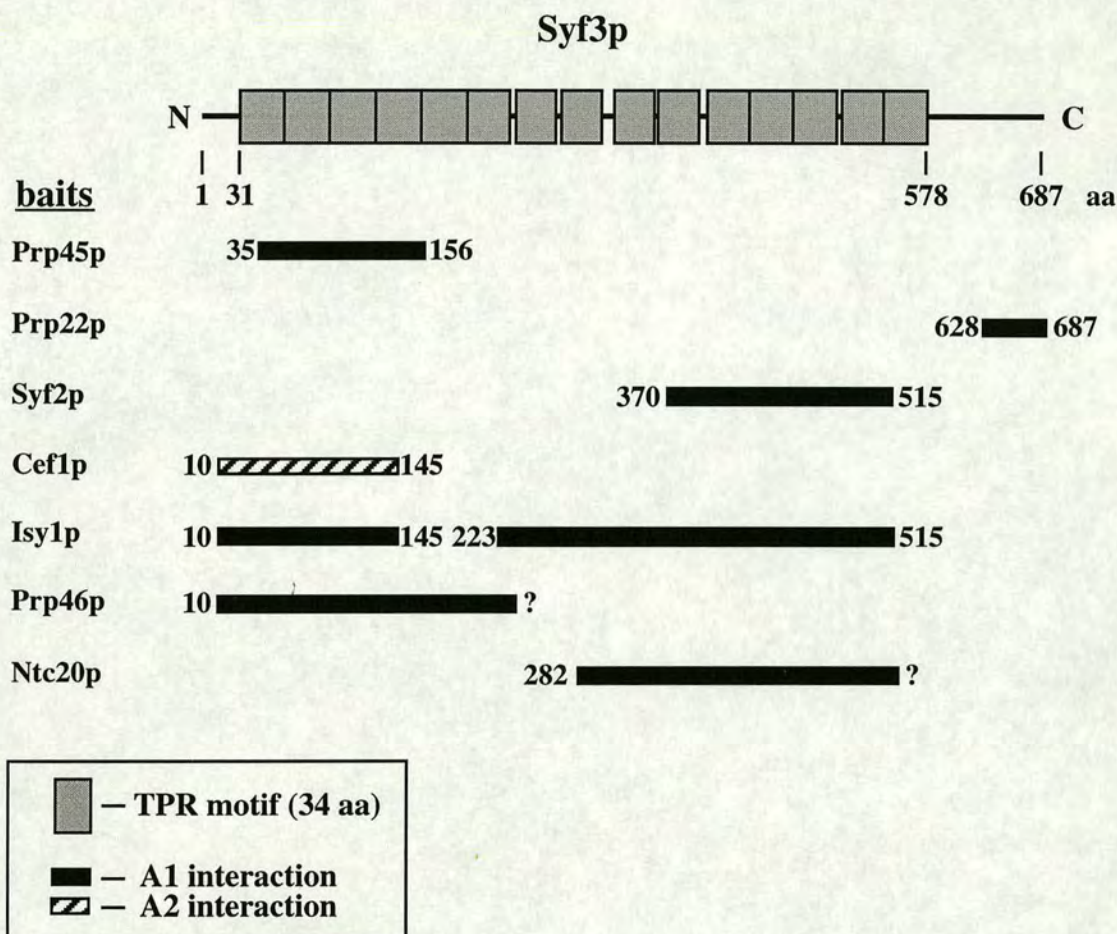
Figure VI.17:
Two-hybrid interactions of the Syf3 protein. Syf3p contains 15 TPR repeat units. The bars underneath the protein indicate the putative Syf3p interacting regions of several proteins used as baits in two-hybrid screens. For A1 interactions the indicated region results from the minimal overlap of the prey fragments isolated. ? - exact end point of the interacting region is not known; the approximate end point was estimated from the length of the insert fragments in the prey plasmid, as determined by restriction digest and electrophoresis. For information on the proteins and references see text.

The only interaction of Syf3p identified so far, which does not involve any of the TPR motifs is that with Prp22p. Prp22p seems to associate with the non-conserved c-terminus of Syf3p. Surprisingly, Chung *et al.,* 1999, demonstrated that the c-terminal third of Syf3p is required neither for splicing nor for cell growth. This could suggest that the observed interaction of the two essential splicing factors is simply an artefact of the two-hybrid system, i.e the interaction is of no biological significance, but occurs due to certain random sequence or structural features of the two proteins that promote an association. However, two arguments speak against this. First, Prp22p was found as a prey with the Syf1p bait, adding some weight to the hypothesis, that the protein is indeed part of the above described network of protein-protein interactions. Second, *PRP22* was identified, like *SYF1, SYF2, SYF3*, in the synthetic lethal screen with the *prp17* deletion mutant (Ben-Yehuda and Kupiec, unpublished). Thus, there is a strong genetic link of *PRP22* to other genes coding for components of this putative protein network, suggesting a related function of the proteins. It could be imagined that the two-hybrid interaction of Prp22p with the dispensable c-terminus of Syf3p is only one half of the truth. Maybe other important binding sites for Prp22p exist within Syf3p, which are simply not accessible, not properly folded or not fully functional in the two-hybrid fusion proteins or protein fragments, to allow recognition by Prp22p.

In conclusion, numerous two-hybrid data suggest the existence of a multi protein complex required for pre-mRNA splicing, containing the newly identified factors Prp45p and Prp46p, which might be a discrete subunit of the spliceosome. It could well be imagined, that the TPR motif proteins Syf3p and maybe Syf1p form a scaffold for the formation of such a complex by providing binding sites for a number of proteins, especially in the amphipatic groove formed by multiple TPR repeats.

VI.11 Discussion

The previous sections described the initial functional characterisation of the Prp45 protein of *S. cerevisiae*. The protein was identified in a two-hybrid screen using the second step- and spliceosome disassembly factor Prp22p as bait. The construction of a strain carrying a conditionally regulated *PRP45* allele allowed the metabolic depletion of the protein from the cells. In the absence of Prp45p, the cells accumulate unspliced precursor RNA, showing that Prp45p is essential for pre-mRNA splicing *in vivo*. In contrast to Prp22p, which is only required for splicing of introns in which the 3' splice site is more than 21 nucleotides away from the branchpoint, Prp45p seems to be required for splicing of all precursor RNAs, since it was demonstrated that splicing of pre-U3 snoRNA, containing an intron with a very short (7 nucleotides) branchpoint-3' splice site distance, was seriously inhibited upon depletion of Prp45p. This suggested a requirement of Prp45p for the general structural formation of the spliceosome and thus for splicing of all pre-mRNAs. It furthermore might indicate that Prp45p does not directly modulate Prp22p function during 3' splice site selection and cleavage, since one would expect a similar intron-dependent requirement for Prp45p as it is exhibited by Prp22p during the second step. Possible is also that the two proteins interact at another stage of the splicing. The finding that, *in vitro*, Prp45p depletion does not lead to a typical second step block with the accumulation of exon 1 and lariat-intron exon 2, but that splicing activity is abolished already before the first transesterification reaction supports this idea. In addition, no other known second step splicing factors were identified as interacting preys in the two-hybrid screen with the Prp45p bait, suggesting that Prp45p is not in proximity to the 3' splice site, when the second transesterification is carried out.

The investigation of the association of Prp45p with the spliceosome revealed that the protein joins the spliceosome before the first transesterification reaction is performed and stays within the complex until after the intron has been removed from the pre-mRNA. Thereafter it can be found associated with the excised intron, but not with the message. Precipitating a protein A-tagged version of Prp45p from whole cell extracts under non-splicing conditions (no added ATP) co-precipitates small amounts of the U2, U5 and U6 snRNAs. The lack of U1 snRNA in the co-precipitates suggests

that the protein is not a tightly associated component of the U1 snRNP, i.e. is unlikely to play a role in the formation of the commitment complex. Since Prp45p strongly co-precipitates pre-mRNA and since, in the absence of Prp45p, splicing *in vitro* is abolished before the first transesterification takes place, it can be supposed that the physical presence and function of Prp45p is required for the assembly of an active spliceosome. The question arises, why the U4 snRNA is not found in the co-precipitates when protein A-tagged Prp45p is immunoprecipitated from cell extracts, although the U4/U6 and U5 snRNPs join the spliceosome as a preformed tri-snRNP? One possibility to explain this, would be to suggest that Prp45p joins the spliceosome just after the U4 snRNP has dissociated and before the active spliceosome performes the first transesterification reaction. However, after addition of the tri-snRNP, the pre-mRNA becomes rapidly processed. It is therefore possible that the U2, U5 and U6 snRNAs in the precipitates are actually derived from post-splicing spliceosomal complexes. Another possibility that could be imagined is, that the n-terminal protein A tag somehow interferes with the association of the U4 snRNP to the pre-spliceosome and therefore, the U4 snRNP is not as efficiently co-precipitated. Maybe it would be more conclusive to deplete residual endogenous ATP from the cell extracts before the immunoprecipitation is done, to ensure that spliceosome formation is prevented. This might give a clearer picture of Prp45p association with spliceosomal subcomponents. Additional clues about the function of Prp45p and the timepoint of its association with the spliceosome can be gained from the data of numerous two-hybrid screens performed in our laboratory, which allow the drawing of an extensive network of protein-protein interactions. Such network is shown in figure VI.18.

Figure VI.18:
Protein-protein interaction network revealed by exhaustive two-hybrid screening of a yeast genomic DNA library.
The two-hybrid screens were performed by ✕ - I. Dix and C. Russell, ◆ - C. Russell and M. Mellor-Clark, ■ - C.Russell and A. Clark; ↗ - A. Colley, ☯ - this work (all this laboratory). For references on the proteins and discussion see text.
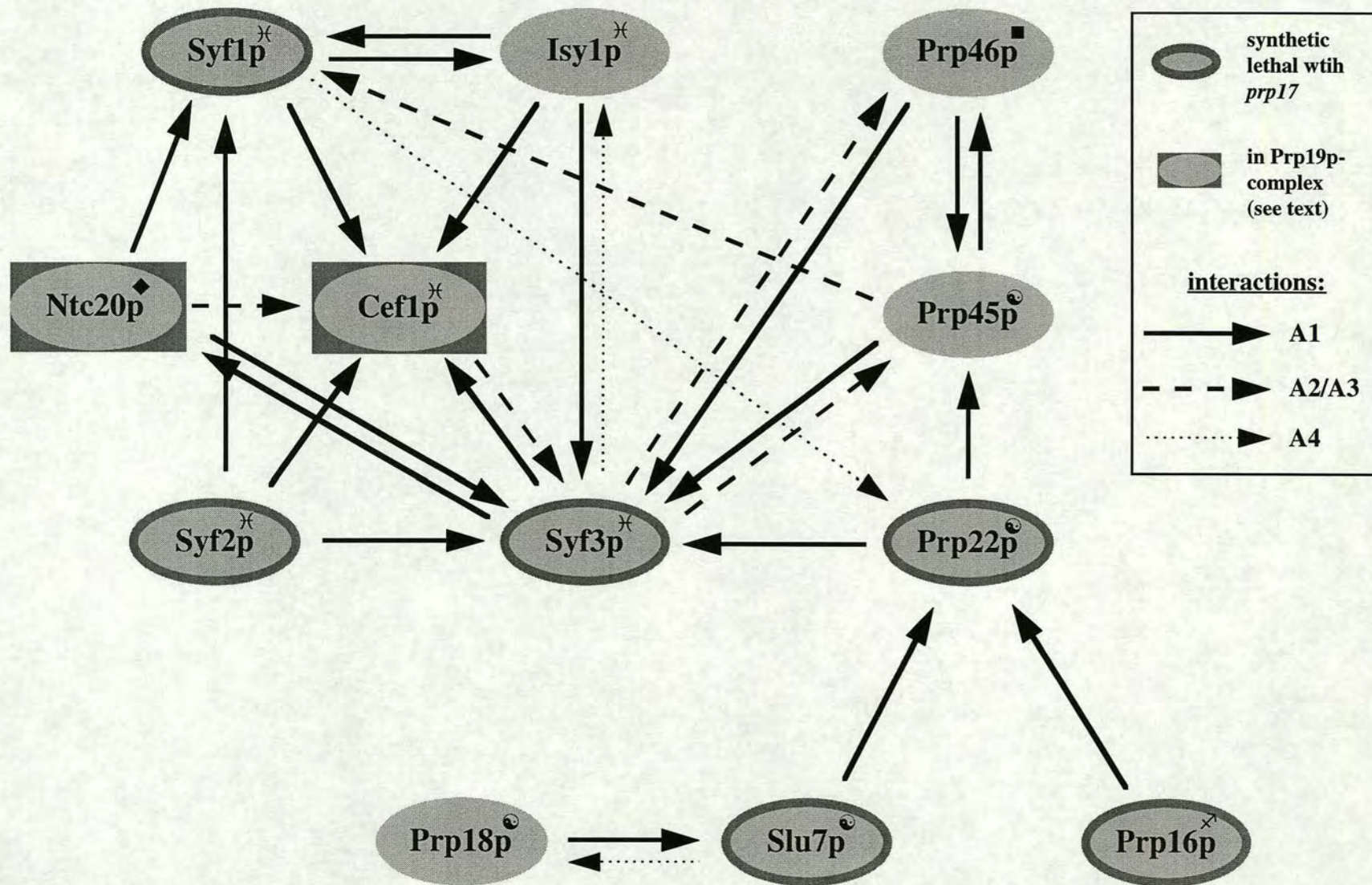
The figure illustrates that via its interaction with Syf3p, Prp45p (as well as Prp22p and Prp46p) is linked with a putative complex of proteins, which again all (apart from Syf1p) interact with Syf3p (see also chapter VI.10.3). The two-hybrid interactions which form the basis of this network are in most cases A1 interactions, they are in many cases found reciprocated and furthermore, the proteins within this network have not been identified in any other two-hybrid screens within the TAPIR network, strongly suggesting that these interactions (or many of them) are indeed specific, probably biologically significant interactions. For some of the proteins, experimental data are available, which could give some indication of a possible role for Prp45p.

As mentioned above, the Ntc20 protein as well as the Cef1 protein are members of the Prp19p associated complex, which contains at least 5 additional proteins (Snt309p, Ntc120p, Ntc90p, Ntc40p and Ntc30p; Tarn *et al.*, 1993 and 1994; Chen *et al.*, 1998; Tsai *et al.*, 1999). For all members of the Prp19p complex so far characterised, it has been shown that they associate with the spliceosome in the same manner, namely concomitantly with or immediately after dissociation of the U4 snRNP, but before the formation of an active spliceosome. This is exactly what could be expected for Prp45p, being a splicing factor required before the first transesterification reaction which does not associate with the U4 snRNP, but with the U2, U5 and U6 snRNPs.

McDonald *et al.* (1999) coprecipitated a high-molecular weight protein complex with the *S. pombe* cdc5 protein, the homologue of the *S. cerevisiae* Cef1p. Interestingly, among the 9 identified proteins were the *S. pombe* homologues of Prp46p, Syf1p, Syf3p and Prp19p, which suggests the existence of related complexes in the two fungi. The authors furthermore showed that the cdc5 protein co-precipitates U2, U5 and U6 snRNAs from cell extracts, but no U4 and U1 snRNAs. Again, this is in agreement with the data available for the *S. cerevisiae* counterparts of the Prp19p complex and with the results yielded for Prp45p. The result of the snRNA co-precipitation with cdc5p, also argues against the hypothesis that the protein A tag could be the reason for the lack of U4 snRNA in the protA::Prp45p co-precipitates. For cdc5p, the co-precipitation of snRNAs was performed with two different tagged versions (cdc5hemaglutinin (HA) and cdc5myc) with the same result. Assuming that Prp45p and Cef1p act at a similar step of the splicing reaction, the same snRNA association pattern could be expected.

The situation does not seem to be exactly the same for all components of the illustrated network. Chung *et al.* (1999) demonstrated a role for the Syf3p (Clf1p) in the assembly of an active spliceosome. However, they show that in extracts depleted of Syf3p, pre-spliceosomes do rapidly form, but that the addition of the tri-snRNP is seriously inhibited, suggesting a requirement for Syf3p for the recruitment of the U5/U4/U6 snRNP to the spliceosome. This would mean that Syf3p acts slightly earlier than other components of the proposed complex, maybe recruiting the tri-snRNP first and then, after dissociation of the U4 snRNP, supporting the association of additional proteins. Considering the extraordinary structure of Syf3p, containing 15 TPR repeat modules, a function in recruiting numerous proteins and organising their interactions could be imagined.

The question remains, why was Prp45p originally identified in the two-hybrid screen with the second step splicing factor Prp22p? Although for Syf1p, Syf3p, Cef1p and Prp45p a requirement for the first step of the splicing reaction has been demonstrated (Chung *et al.*, 1999; Tsai *et al.*, 1999; Russell, unpublished results and this work), it has to be noticed that three of the proteins in the above described putative complex indeed have been genetically linked to the second step splicing factor Prp17p (see also figure VI.18). *SYF1*, *SYF2* and *SYF3* (as well as the genes for the second step splicing factors *PRP22*, *SLU7* and *PRP16*) were identified in the aforementioned synthetic lethal screen with *PRP17* (Ben-Yehuda *et al.*, unpublished results). It is therefore not unlikely, that Syf1p, Syf2p, Syf3p and Prp45p have additional functions during the second step. All four proteins remain attached to the spliceosome until the second step is completed (Russell, unpublished data; this work) and therefore could be involved in the second catalytic step or even at a stage when Prp22p acts in releasing the mature message from the spliceosome.

Chapter VII

Prp46p is required for pre-mRNA splicing in vivo

VII.1 Introduction

The Prp46 protein was identified as the strongest interacting prey protein in a two-hybrid screen with the novel splicing factor Prp45p. A region containing 7 WD repeat units within Prp46p seems to be required for the interaction with Prp45p. This chapter will discuss putative homologues of Prp46p and what is known about their function. In addition, evidence will be provided that Prp46p is a splicing factor of *S. cerevisiae*.

VII.2 *PRP46* encodes a conserved WD protein

Primary structure analysis of the 451 amino acid Prp46 protein revealed that it is a member of the ancient regulatory family of WD proteins (for a review on WD proteins see Smith *et al.*, 1999). WD proteins are made up of highly conserved repeating units which usually end with the amino acid residues tryptophan and aspartatic acid (WD). A single protein can contain 4-16 of the WD modules. The proteins are present in all eucaryotes, but are very rare in prokaryotes. Despite their high degree of conservation and modular organisation, which also predicts a conserved three-dimensional structure, WD proteins are of diverse functionality and serve in numerous cellular processes. Several members have been shown to function in signal transduction pathways, such as the Gβ subunit of the heterotrimeric G proteins, well characterised components of the transmembrane signalling machinery (Wall *et al.*, 1995). Others are transcriptional regulators, for example the TFIID subunit of the TATA-box binding complex (Hoey *et al.*, 1993). A number of WD proteins, such as Cdc40p and Cdc20p, are important for cell cycle progression (Farruggio *et al.*, 1999; Ben-Yehuda *et al.*, 1998), yet others regulate vesicle formation and vesicular trafficking (Pryer *et al.*, 1993). Other proteins have been found to be involved in sulfur metabolism in fungi (Natorrf *et al.*, 1998). Among the members playing a role in RNA processing are, for example, the mammalian cleavage-stimulation factor associated protein CstF, required for polyadenylation (Takagaki and Manley, 1992) or the yeast Mak11p, needed for double stranded RNA replication (Icho and Wickner, 1988).

So far, two WD proteins have been shown to be required for pre-mRNA splicing in *S. cerevisiae*: the second step splicing factor Prp17p/Cdc40p (Jones *et al.,* 1995) and the Prp4 protein, which acts at late stages of spliceosome assembly (Ayadi *et al.,* 1997). Both proteins have orthologs in human (Ben Yehuda *et al.,* 1998; Wang *et al.,* 1997; Zhou and Reed, 1998).

Besides the general degree of sequence conservation with other WD proteins (usually between 15 and 20% amino acid identity over the entire polypeptide), Prp46p shares a much stronger homology to a subset of proteins within the family, which are therefore likely to be functional equivalents: these proteins include a human protein (Acc.no. AF044333) of 494 amino acids (aa) with a predicted molecular weight of 57 kDa, a hypothetical 513 aa mouse protein (AF044334) , the 473 aa long prp5 protein of *S. pombe* (O13615), a hypothetical 494 aa long protein of *C. elegans* (CAA98274) and the pleiotropic regulatory proteins PRL1 and PRL2 of *A. thaliana* (486 and 479 aa). A primary structure alignment (GeneStream Align) reveals an amino acid identity of the full length Prp46p to the human homologue of 38.9%, to the *S. pombe* prp5p of 44.6%, to the *C. elegans* protein of 38.9% and to the plant PRL1 and PRL2 proteins of 41.2% and 40%, respectively.

(A complete alignment of these proteins was made by Nemeth *et al.,* 1998). All proteins contain a c-terminally located cluster of seven WD repeat units, with no spacing between the repeats. The sequence of Prp46p, with the WD-repeats aligned, is shown in figure VII.1.

Although the homology of Prp46p to the other proteins is highest in the WD repeat cluster (amino acids 122-427; 50-59% amino acid identity), there is significant homology also at the n-terminus (aa 1-121; 18-21%) and at the small c-terminal extension (aa 428-451; 37-50%), underlining the hypothesis that these proteins represent functional homologues.

1   MDGNDHKVENLGDVDKFYSRIRWNNQFSYMATLPPHLQSEMEGQKSLLMR

51 YDTYRKESSSFSGEGKKVTLQHVPTDFSEASQAVISKKDHDTHASAFVNK

101 IFQPEVAEELIVNRYEKLLSQ

	variable region	constant region		
122	RPEWHAPWK**L**SRV**IN**GH**LGW**VR**CVAI**D**PVDN**EWF**ITGS**N**D**-TTMKV**WD**
169	**LATGKLKTTLA**GH**VMTVRDVAVS**DRHP**-YLF**SVS**ED**-KTVKC**WD**
211	**LEKNQIIRDYY**GH**LSGVRTVSIH**PTLD**-LIATAGRD**-SVIKL**WD**
253	MR**TRIPVITLV**GH**KGPINQVQCT**PVDP**Q-VVS**SST**D**-ATVRL**WD**
295	VVAGKTMKV**LT**HH**KRS**VR**ATALH**PKEF**S-**VAS**ACT**D**--DIR**S**WG**
336	**LAEGSLLTNF**E**SE**KTG**IINTLSI**NQDD**--VLFAGGD**NGVLSF**YD**
378	YKSGHKYQS**LATREMVGS**L-EG**ERSVLCSTFD**KTGL**R-LITGEAD**-KSIKI**WK**

428 QDETATKESEPGLAWNPNLSAKRF   451

Figure VII.1:
Amino acid sequence of the Prp46 protein. The seven WD repeats (amino acids 122-427) are depicted aligned and the positions of the WD repeat variable and constant regions a are indicated. Conserved residues are highlighted in bold. Amino acid residues on positions which define the regular expression pattern of the WD motif defined by Neer *et al.* (1994) are indicated in vertical boxes (see text).

Intriguingly, the Prp46p homologues from human and fission yeast have been implicated in pre-mRNA splicing: As well as the Prp45 protein homologue Snw1 (or Skip), as reported in chapter VI.2, Neubauer *et al.* (1998) co-precipitated the human Prp46p homologue with splicing complexes from HeLa nuclear extracts. In addition, they showed the *in vivo* co-localisation of GFP-tagged human Prp46p with splicing snRNPs by confocal fluorescence microscopy, confirming the precipitation result. The *S. pombe* prp5 protein was identified through isolation of the temperature-sensitive *prp5-1* mutation which causes a defect in pre-mRNA splicing at the restrictive temperature *in vivo* (Potashkin *et al.*, 1998). The mutant strain also exhibited a cell cycle defect. As shown by fluorescence-activated cell sorting (FACS) analysis, *prp5-1* mutant cells are stalled after DNA-replication, suggesting that they proceed normally through S phase at the restrictive temperature and thus, that the cell cycle is probably arrested in G2-phase prior to mitosis. Soon after, it was shown that the prp5 protein is

present in the aforementioned (chapter VI.2) high molecular weight complex of splicing and cell cycle proteins, which was co-precipitated with the cdc5 protein, the homologue of the *S. cerevisiae* Cef1p splicing factor (McDonald *et al.*, 1999). Besides the observed physical interaction (the authors also show the reciprocal precipitation of cdc5p with prp5p in cell extracts), *prp5* and *cdc5* interact genetically with each other: a *cdc5-120/prp5-1* double-mutant strain exhibits a reduced restrictive temperature compared to either of the single mutant strains.

The biochemical function of the *A. thaliana* proteins PRL1 and PRL2 is less clear. However, as the name suggests (PRL=pleiotropic regulatory locus), a recessive mutation (*prl-1*) that leads to truncation of the c-terminal 94 amino acids of the PRL1 protein, induces a variety of defects in the plant, possibly due to an altered control of glucose and hormone responses (Nemeth *et al.*, 1998). The authors show that the *prl-1* mutation results in changed carbon partitioning patterns and hypersensitivity of the plant to glucose and sucrose. High concentrations of external glucose or sucrose resulted in a reduction of root elongation and a decrease of shoot development as well as a stimulated accumulation of sugars and starch in the leaves. Furthermore hypocotyl elongation slowed down and a premature initiation of side roots was observed. The latter effect suggested an enhanced sensitivity of the seedling to auxins. Indeed, further investigation confirmed hypersensitivity of the plant to a number of growth hormones in addition to auxin such as cytokinin, ethylene and abscisic acid. The authors explain the observed phenotype by proposing a cellular function of PRL1 as a suppressor of glucose responsive genes. Since there is a complex crosstalk between hormonal and metabolic regulation and because light signalling is modulated by glucose and cytokinin, and vice a versa, glucose and cytokinin metabolism is controlled by light, mutations relieving glucose repression are therefore expected to result in defects in cytokinin signalling, root development, general stress responses and chlorophyll and anthocyanin biosynthesis (for a review see Smeekens and Rook, 1997).

The initial molecular characterisation of PRL1 led to the identification of two potential interacting proteins. Nemeth and co-workers noticed the presence of a sequence motif within WD repeats 3 and 6 of PRL1, that shared similarity to a sequence mediating the interaction of RACK1 receptor (another WD protein) with activated protein kinase C (PKC) in mammals (Ron *et al.*, 1994). In *in vitro* binding experiments using a

glutathione-S-transferase (GST) fusion protein of PRL1, indeed PRL1, like RACK1, bound specifically to PKC-βII, but not to PKC-βI or PKC-γ. The data also indicated that the interaction was mediated via the c-terminal 52 amino acids of PKC-βII, since apart from these 52 amino acids PKC-βII and PKC-βI are identical. Notably, the similarity of PRL1 (and therefore Prp46p) with RACK1 is very significant, going beyond the general similarities of WD proteins: first, the modular organisation of seven WD repeats, with only very little spacing in between the repeats is conserved. Second, the seven repeats are organised in such a way, that 4 repeats, which fit 100% into the regular expression pattern of a WD repeat, are followed by 3 WD repeats of relatively degenerate sequence homology with respect to the regular expression pattern. The regular expression pattern (as defined by Neer *et al.*, 1994) was deduced from an alignment of hundreds of WD repeats, and describes which amino acids are most likely to be situated at certain conserved positions within a repeat. Figure VII.1 illustrates in an alignment of the WD repeats of Prp46p, to what extent the amino acid sequence matches this regular expression pattern. It is supposed that a similar arrangement of WD modules (like for Prp46p and RACK1) reflects to a similar biochemical function.

A two-hybrid screen of an *Arabidopsis* cDNA expression library using the full length PRL1 protein as a bait revealed a strong interaction with ATHKAP2 (Acc. no. Y09511), which shows remarkably high sequence identities to α-importins, such as ATHKAP1 of *Arabidopsis*, human HSRP1, Xenopus IMP1, yeast Srp1p, involved in nuclear import of proteins and RNAs. Subsequent *in vitro* binding studies allowed mapping of the interacting domain to a carboxy terminal region of 74 amino acids, which includes the last WD-repeat and a putative SV40-type nuclear localisation signal. Indeed, it was demonstrated that PRL1 localises to the nucleus in *Arabidopsis* cells as well as in green-monkey COS-1 cells, illustrating also the remarkable conservation of the protein in different organisms.

It remains to be determined to what extent the high degree of sequence homology of this novel subset of WD proteins reflects an involvement of the proteins in the same cellular process and a functional analogy. The fact that both the human and the *S. pombe* proteins are involved in pre-mRNA splicing suggests that the other candidates are also splicing factors. For the *A. thaliana* PRL1 protein it remains to be determined

whether the effects of the *prl-1* mutation seen on glucose metabolism and hormone regulation might be secondary effects, which originate in a malfunction of the splicing apparatus.

The observed two-hybrid interaction with Prp45p adds even more weight to the hypothesis that the *S. cerevisiae* Prp46 protein is required for splicing. The next sections present its initial functional characterisation.

VII.3 *PRP46* is an essential gene

In order to approach a functional analysis of Prp46p, a gene deletion experiment was performed. Using oligonucleotides PLKO-1 and PLKO-2 the *HIS3* gene was PCR amplified from plasmid YIp1. After purification of the PCR product, it was transformed into the diploid yeast strain BMA38. The primers contained flanking sequences homologous to regions just upstream and downstream of the *PRP46* ORF, thus allowing replacement of the ORF by the *HIS3* gene locus due to homologous recombination of the PCR product into the chromosome. Transformants were selected for histidine prototrophy and integration of the disruption cassette was checked by PCR using oligonucleotides PLKO-1 and 151-PR2, which is complementary to a region approximately 300 bp downstream of the *PRP46* ORF. A single product of the expected size of 1.4 kb was amplified from the transformants, but not from the wt strain suggesting integration at the desired locus (figure VII.2A). The PCR product was gel purified and sequenced using primer 151-PR2, which confirmed the replacement of the *PRP46* ORF by the marker gene.

The diploid strain (designated YMA151KO1) was subsequently sporulated and the resulting tetrads were dissected onto YPD plates (figure VII.2B).
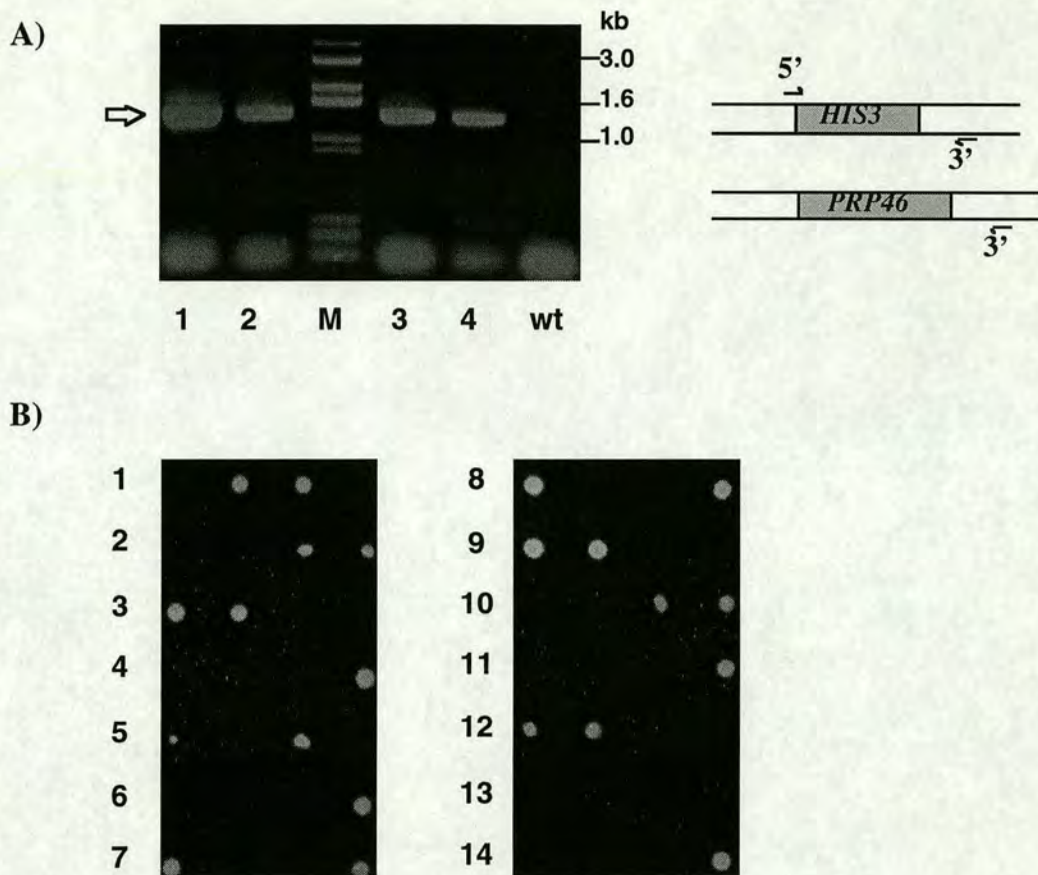
Figure VII.2:
Deletion of the *PRP46* ORF from the genome.
A) PCR on yeast transformants to test for integration of the *HIS3* marker gene into the chromosomal *YMR44w* locus. After transformation of the BMA38 strain with a linear PCR product containing the *HIS3* marker gene flanked by approximately 40 base pairs of the *PRP46* locus at either site, histidine prototrophs were streaked out onto fresh YMM -H medium for colony purification. Then a single large colony of the transformants and of the wt parental strain BMA38 was suspended in 0.02N NaOH, boiled for 5 minutes and an aliquot used in a PCR (II.3.2.8.2) using oligonucleotide primers PLKO-1 (5') and 151-PR2 (3'). The positions on the template at which the primers anneal are indicated in the upper schematic drawing of the *PRP46* locus (either wild-type or after *HIS3*-integration). Aliquots of the PCR reactions were run on a 1% (w/v) agarose gel. The white arrow indicates the position of the amplified 1.4 kb fragment, which was expected in case of successful integration of the *HIS3* cassette into the *PRP46* target locus. **B)** Tetrad analysis of strain YMA151KO1. Strain YMA151KO1was grown overnight at 30°C in YPDA, the cells collected and transferred to SPM sporulation medium. After 5 days of incubation at 23°C the formed tetrads were dissected onto YPD agar plates. The spores were incubated for 3 days at 23°C.

For all of the 14 tetrads dissected, at most two of the spores germinated and grew to colonies, demonstrating that *PRP46* encodes an essential gene. All colonies grown were shown to be histidine auxotrophs, confirming that the lethal phenotype segregated with the *HIS3* marker and thus is due to replacement of the *PRP46* ORF (data not shown).
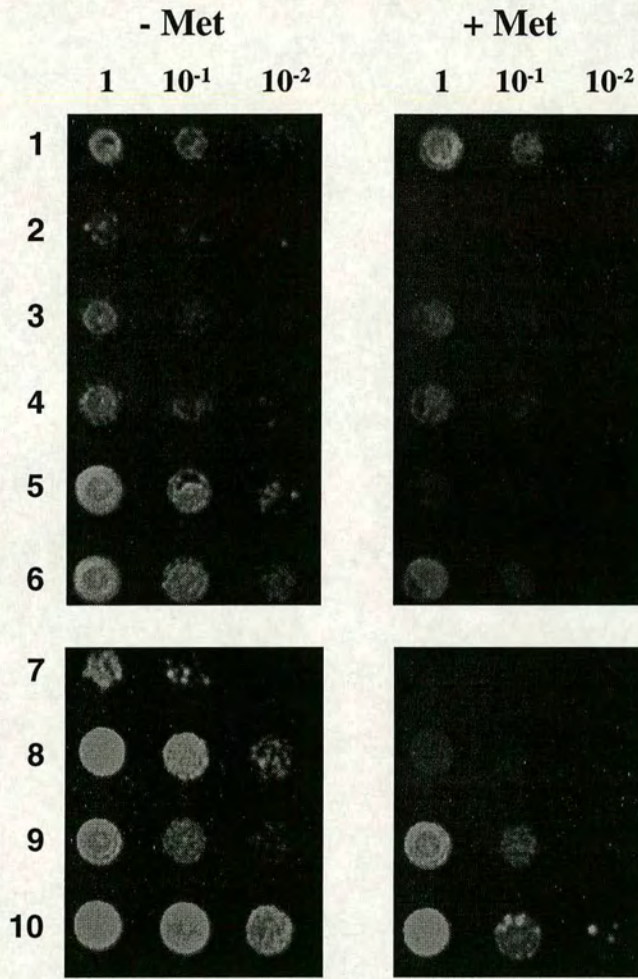
VII.4 Generation of a conditionally regulated and epitope-tagged *PRP46* allele

In order to allow the initial functional characterisation of the essential Prp46p, a conditionally regulated allele of the gene was generated. Again the method developed by Lafontaine and Tollervey (1996) was used to integrate a cassette, containing a reporter gene, a regulatable promotor and a sequence encoding an epitope tag upstream the target ORF. Replacement of the native promotor region with such a cassette should allow repression or activation of transcription of the target gene depending on the choice of growth medium. In this case, a cassette was PCR-amplified from vector pUC19-55HA$_2$ (constructed by R. van Nues, this laboratory), using oligonucleotides 151MetA and 151MetB. This cassette contained the *TRP1* marker gene, the promotor region of the *MET3* gene, which can be repressed by the presence of methionine in the growth medium, as well as a sequence encoding a double hemagglutinin A (HA) epitope tag. The PCR-product was gel-purified and subsequently transformed into the haploid strain BMA64n, which is deleted for the *TRP1* gene. Tryptophan prototrophic transformants were picked and streaked once to yield purified clones. The colonies were then suspended in microtiter plates and dilutions transferred onto YMM -W plates, either without or containing 7 mM methionine (figure VII.3A). Although not in all cases, some clones indeed showed highly reduced growth on methionine-containing medium, suggesting that *PRP46* transcription was repressed, and that Prp46p was depleted from the cells.
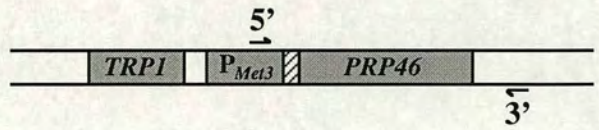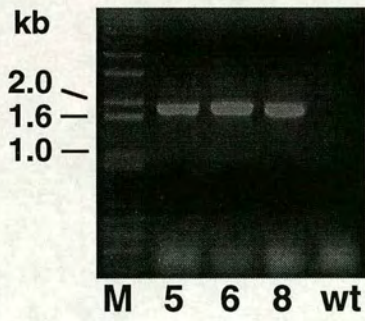
Figure VII.3:
Conditional regulation of *PRP46* on the chromosome.
Yeast strain BMA64n was transformed with a linear DNA fragment (generated by PCR- amplification from vector pUC19-55HA$_2$) consisting of the *TRP1* marker gene, the *MET3* promotor region (P_{Met3}) and a double HA-epitope tag (HA$_2$), flanked by approximately 40 base pairs homologous to sequences upstream of the *PRP46* locus. **A)** Transformants were selected on YMM -MW medium and streaked for colony purification. Single colonies were then suspended and serially diluted into microtiter plates and spotted onto YMM -M medium (-Met) or YMMsup medium containing 7 mM methionine (+ Met). The plates were incubated for 2 days at 30°C. **B)** Yeast colony PCR on cells from clones 5, 6 and 8 (picture A) using oligonucleotide primers Met3 (5') and 151-PR2 (3'). The schematic drawing depicts the *PRP46* locus after integration of the regulation cassette. The annealing positions for the primers are shown. The dashed box indicates the position of the sequence encoding the double HA tag.

A)

- Met   + Met

1   10⁻¹   10⁻²   1   10⁻¹   10⁻²

1
2
3
4
5
6

7
8
9
10

B)

kb

2.0
1.6
1.0

M   5   6   8   wt

5'

TRP1   P_Met3   PRP46

3'

To prove that integration had happened at the target locus, oligonucleotides Met3 and 151-PR2 were used in a yeast colony PCR to amplify a region of the *PRP46* locus after integration of the cassette (figure VII.3B). As shown for clones 5, 6 and 8, a PCR product of approximately 1.8 kb was obtained, which could not be amplified in wt cells, suggesting correct integration. The PCR product of clone 5 was subsequently gel-purified and sequenced using oligonucleotide Met3. Indeed, the sequence analysis confirmed the integration of the cassette upstream the *PRP46* ORF (data not shown). The strain containing the regulatable *PRP46* gene will henceforth be referred to as YMA151/2.

The effect of *PRP46* repression was then studied for growth in liquid medium either lacking or containing 7 mM methionine. A wt culture grown in the presence of 7 mM methionine served as control. Samples of the cultures were taken at different time points and the optical density (OD_{600}) of the cultures was measured. The resulting growth curve is depicted in figure VII.4. It can be seen in medium without methionine, that strain YMA151/2 grows essentially like the wt strain (the growth rate of the wt strain BMA64n is hardly influenced by the addition of 7 mM methionine, data not shown). However, the presence of 7 mM methionine in the medium increases the doubling time of strain YMA151/2 from 3.1 to 5.5 hours. This suggests that *PRP46* expression is repressed to some extent upon addition of methionine, but that the repression is not very tight. The essential Prp46p probably does not become completely depleted from the cells and therefore cell growth is only slowed down, but not abolished. Even extended growth (24 hours) in medium containing 20 mM methionine does not lead to a further reduction of the growth rate (data not shown). The slow growth phenotype of strain YMA151/2 in methionine containing medium suggests that even partial depletion of Prp46p has a significant effect on the metabolic pathway in which it functions. Therefore, it was tested whether Prp46p depletion leads to a recognisable splicing defect, because the two-hybrid data obtained in this work, as well as the available data from putative homologues in other organisms suggested a role of the protein in the splicing process.
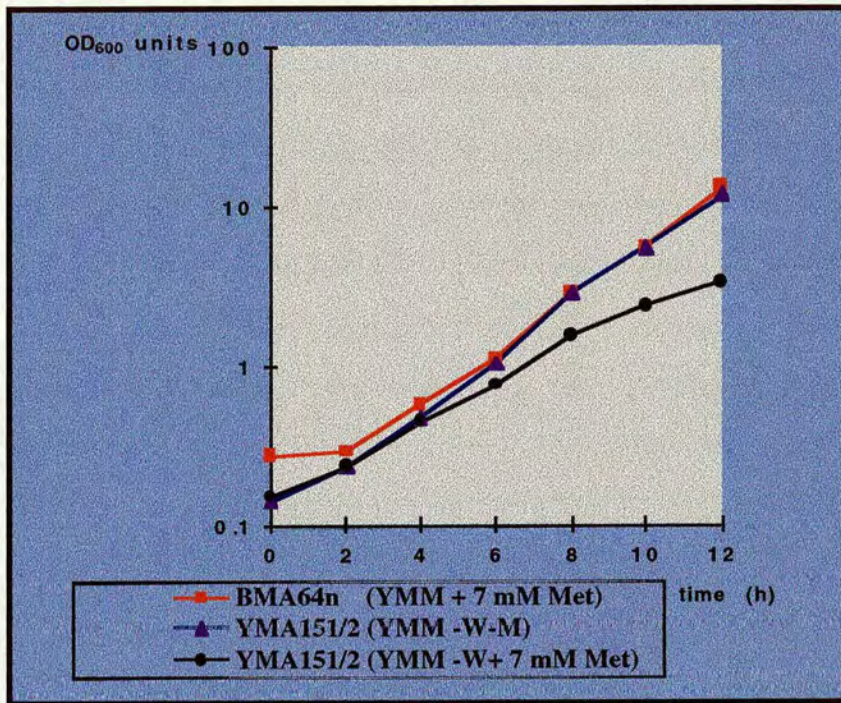
Figure VII.4:
Effect of Prp46p depletion on cell growth.
Overnight cultures of strains YMA151/2 and BMA64n were grown at 30°C in YMM -WM or YMMsup liquid medium to midlogarithmic phase and aliquots of the cultures were used to inoculate pre-warmed 250 ml of either YMMsup + 7 mM methionine (for BMA64) or YMM -WM and YMM -W + 7 mM methionine (for YMA151/2) to an OD_{600} of 0.15-0.35. The cultures were grown at 30°C and aliquots were removed in order to monitor growth by measuring the OD_{600} at different time points. The cultures were diluted at intervals to maintain logarithmic growth.

VII.5 Prp46p is required for pre-mRNA splicing *in vivo*

Cells from strain YMA151/2 were grown in YMM -W liquid medium either with or without 7 mM methionine. As control, the parental wt strain BMA64n was grown in YMMsup liquid medium containing 7 mM methionine. At different time points samples of the cultures were taken and total RNA was prepared. The processing of the precursor of U3 snoRNA was assayed by primer extension analysis. Levels of U1 snRNA were also investigated as loading control. The results of this primer extension assay are shown in figure VII.5.

Already after 4 hours growth under repressing conditions, clear bands corresponding to the U3A and U3B precursor molecules were detectable (lane 2). These bands did not appear when cells were grown under permissive conditions (lane 1). This strongly suggests that repressing *PRP46* expression, thereby depleting the protein from the cells, leads to a severe splicing defect that causes the accumulation of unspliced precursor RNAs. The accumulation is even more significant at later time points. The presence of methionine in the medium does not effect the processing of the U3 snoRNA in the wt strain (lane 3 at 4, 8 and 12 hour time points), confirming that the observed splicing defect in YMA151/2 is indeed due to the repressing effect of methionine on the *MET3* promotor and not to a general toxicity of high concentrations of methionine. Therefore, it can be concluded that Prp46p is required for pre-mRNA splicing *in vivo*, and thus, that Prp46p represents a novel splicing factor of *S. cerevisiae*.
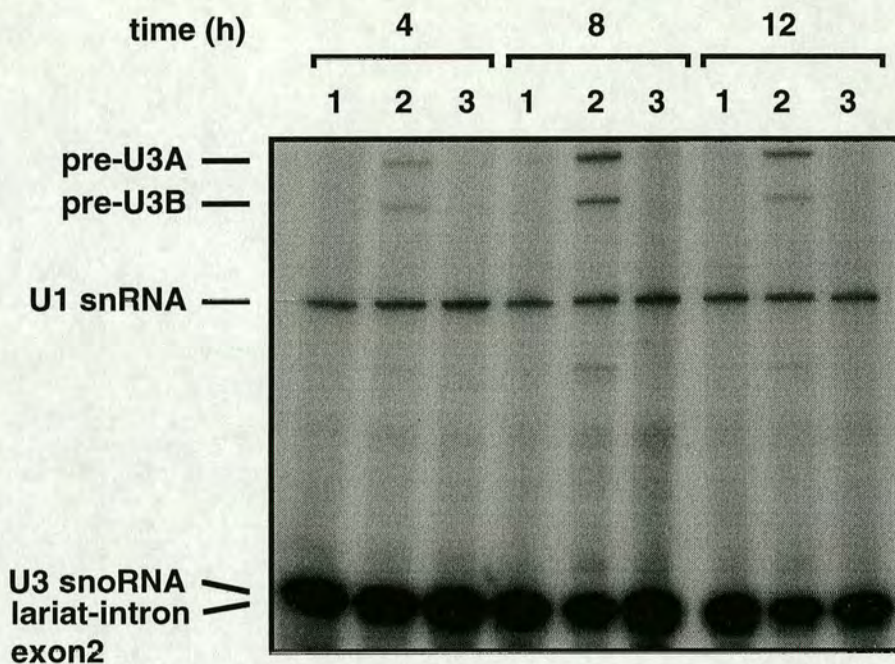
Figure VII.5:
Primer extension analysis of pre-mRNA splicing in strain YMA151/2.
Cultures of YMA151/2 and BMA64n were grown as described in figure legend VII.4
and 10 ml samples were taken at different time points. Total RNA was extracted from
these samples and 10 µg of RNA was used in a primer extension reaction using
radiolabelled oligonucleotide primers complementary to the extreme 5' end of exon 2
of the U3 snoRNA and to the (intronless) U1 snRNA. The reactions were
deproteinised, extracted with phenol/chloroform and the products resolved on a 6%
(w/v) polyacrylamide gel. The result was visualised by autoradiography. The
positions of the primer extension products are indicated; time (h) - time point of
sample substraction; 1 - YMA151/2 cells grown in YMM -WM, 2 - YMA151/2 cells
grown in YMM -W + 7 mM methionine, 3 - BMA64n cells grown in YMMsup + 7
mM methionine.

VII. Discussion

The previous sections describe the characterisation of Prp46p, which was originally isolated in a two-hybrid screen with the Prp45p bait. It is demonstrated, that Prp46p is essential for cell viability. A strain was generated in which *PRP46* expression could be conditionally regulated, which allowed (at least partial) depletion of Prp46p from the cells. Using this strain, it was demonstrated that Prp46p depletion leads to a splicing defect *in vivo*. Therefore, Prp46p is a bona fide pre-mRNA splicing factor of *S. cerevisiae*.

The extraordinary and conserved primary structures of Prp45p and the WD protein Prp46p, as well as information that is available for their homologues, leave some room for speculation with respect to Prp45p and Prp46p function. Interestingly, for the *A. thaliana* homologue of Prp46p a specific interaction with protein kinase C-βII was observed (Nemeth *et al.* 1998). Indeed, in a Prp46p two-hybrid screen, an A3 interaction with a yeast Pkc1 protein (ORF *YBL105c*) has been detected (C. Russell and A. Clark, unpublished). In addition, Prp45p and Cef1p baits interacted with Pkc1p in the two-hybrid screens (A3 and A1 interactions; Dix and Russell, unpublished; this work). These findings suggest that phosphorylation events might influence the interactions. Another finding in this context is intriguing: the mammalian RACK1 protein, which is a receptor for activated protein kinase C, has been shown to inhibit the tyrosine kinase activity of src tyrosine kinases (Chang *et al.*, 1998). As mentioned above, RACK1 represents a WD protein which can be grouped into the same structural subgroup of WD proteins as Prp46p and its homologues. Prp45p, in turn, and in particular its homologues from higher eukaryotes, show a significant similarity to src tyrosine kinases, although it remains unclear whether they in fact have kinase activity: the proteins contain the n-terminal half of an SH2 domain (although to some extent degenerate), they (apart from Prp45p) have a conserved proline rich region, maybe representing an SH3 domain binding site and furthermore, they contain three absolutely conserved tyrosine residues, which could be potential targets for phosphorylation. Thus, it could be imagined that, just as activated PKC-βII interacts with RACK1 which then in turn regulates the activity of src tyrosine kinases, Pkc1p interacts with Prp46p, which then modulates Prp45p function. Interestingly, a

regulation of pre-mRNA splicing by the tyrosine kinase activity of src has been demonstrated (Neel *et al.*, 1995). By overexpressing activated *src*, a significant reduction of splicing activity was observed. Partially spliced reporter constructs accumulated not only in the nucleus, but also in the cytoplasm, suggesting that also the export of incompletely spliced precursors was effected. Moreover it was demonstrated, that the regulation of transcription and pre-mRNA processing by src, can be uncoupled. Different domains of src are important for distinct processes, suggesting that different effector pathways might be involved, which regulate general transcription, pre-mRNA splicing or mRNA transport (Gondran and Dautry, 1999).

A regulation of splicing activity via signal transduction cascades that involve the homologues of Prp45p and Prp46p in higher eucaryotes is therefore a possibility. Whether Prp45p, Prp46p or other splicing factors in yeast would also act via versatile mechanisms like this remains to be determined.

Since Prp46p was shown to interact in a strong and reciprocal way with Prp45p as well as with Syf3p in the two-hybrid screens, it is likely to be functionally linked or an integral part of the putative protein complex that is discussed in some detail in chapter VI.11 (see also figure VI.18). Furthermore, the *S. pombe* homologue of Prp46p (prp5) was shown to be a component of a related complex which was precipitated together with the cdc5 protein (Cef1p in *S. cerevisiae*), supporting the idea of the presence of Prp46p in this putative complex of proteins (McDonald *et al.*, 1999).

Interestingly, it was shown that a mutation in the *S. pombe prp5⁺* gene (*prp5-1*) not only leads to a splicing defect, but also to a clear block in progression of the cell cycle from G2 phase into mitosis (Potashkin *et al.*, 1998). Cell cycle defects have been demonstrated for many of the proteins in the related complexes in *S. pombe* as well as in *S. cerevisiae* . In *S. pombe*, apart from prp5p, also cdc5p function is required for G2-mitosis transition (Ohi *et al.*, 1998). In *S. cerevisae*, mutations in the *SYF1* and *SYF3* as well as in the *CEF1* gene cause the cells to arrest in G2 phase (Ben-Yehuda *et al.*, unpublished; Ohi *et al.*, 1998). Furthermore, a requirement of the *S. cerevisiae* Prp22p and a homologous protein in *S. pombe*, cdc28, for the progression of the cell cycle from G2 to mitosis has been demonstrated (Lundgren *et al.*, 1996; Hwang and

Murray, 1997). In the synthetic lethal screen with the *prp17/cdc40* deletion mutant, which led to the identification of the three *SYF* genes, mutant alleles of the second step splicing factors *slu7*, *prp16* and *prp8* were isolated (Ben-Yehuda *et al.*, unpublished). For *PRP8* is was shown earlier, that a mutant allele, *dbf3-1* resembled the cell-cycle defect phenotype of *prp17/cdc40* strains, in which DNA replication was delayed and cells arrested in G2 at the restrictive temperature (Shea *et al.*, 1994).

The finding that numerous splicing factors are also required for cell cycle progression raises the question, whether simply one or more intron-containing genes become rate limiting for cell cycle progression, or whether these processes are linked in a regulatory manner? Different mechanisms could be imagined, to explain how cell cycle progression is effected by the splicing machinery, or vice versa: Maybe some defects within the splicing machinery are sensed by a cell cycle checkpoint and a cell cycle arrest results. Alternatively, the removal of certain introns at a specific stage of the cell cycle leads to the activation of a specific splicing complex, such as the one described in the previous sections. Since numerous protein kinases are involved in the regulation of the cell cycle, phosphorylation and dephosphorylation could play a major role in activating such a splicing complex. Prp45p and Prp46p, being plausible targets for regulation via phosphorylation, as discussed above, might therefore be candidates which connect the cell cycle with the splicing machinery in a regulatory manner.

Chapter VIII

Overview and future work

In this work Prp45p and Prp46p, two novel pre-mRNA splicing factors in *Saccharomyces cerevisiae* were identified through exhaustive two-hybrid screens. Both proteins are evolutionarily conserved and essential for cell viability, which emphasises their requirement for constitutive cell metabolism. Prp45p is associated with the spliceosome throughout the splicing reactions and is required already prior to the first transesterification reaction. This suggests that either it is simply the physical presence of the protein that is essential for a proper structural formation of an active spliceosome or that is has regulatory functions necessary for the catalytic activity of the spliceosome in step 1. Numerous two-hybrid screens performed in this laboratory revealed a network of protein-protein interactions (see figure VI.18), which suggests the existence of a putative protein complex, of which Prp45p and Prp46p might be components, or at least, with which they associate at one point or another.

For some of the proteins within this network it had already been determined, at what stage they join the spliceosome and what their functions might be: Syf3p (Clf1p) seems to be required for the association of the tri-snRNP onto the pre-spliceosome (Chung *et al.*, 1999), Ntc20p and Cef1p join the spliceosome concomitantly with or just after dissociation of the U4 snRNA , i.e. slightly later than Syf3p (Tsai *et al.*, 1999). Chung *et al.* (1999) furthermore found that Syf3p interacted with components of the cross-intron bridge, Mud2p and Prp40p (see figure I.4), in directed two-hybrid assays and proposed a model for Syf3p action, in which it joins the spliceosome at the same time as (or even bound to) the U2 snRNP, thereby displacing the branchpoint bridging protein BBP from the pre-mRNA. This rearrangement then allows addition of the tri-snRNP. Indeed, Rutz and Seraphin (1999) showed recently that BBP cannot be found in pre-spliceosomes, so it must leave upon U2 snRNP addition. Prp45p, being a strong and reciprocal two-hybrid interactor with Syf3p is a good candidate to be involved in promoting these rearrangements. Interestingly, Fromont-Racine *et al.* (1997) identified Prp22p as A3 prey in a two-hybrid screen with BBP as bait. This is intriguing, since in the work presented here, Prp22p is shown to interact with Prp45p as well as Syf3p. Thus, there are accumulating data, which suggest that Prp22p and Prp45p are involved in promoting the replacement of BBP upon Syf3p (and U2 snRNP) addition to the commitment complex. It would be interesting to investigate the effect of Prp45p-depletion on spliceosome formation by resolving splicing complexes from Prp45p-depleted extracts

by non-denaturing gel electrophoresis. From the above model, one would expect that pre-spliceosomes do form, but that, like in Syf3p-depleted extracts, tri-snRNP addition is affected or prevented. Another point, that would be worth studying as a consequence of the above considerations is whether Prp45p does bind to the substrate pre-mRNA. If the RNA-binding BBP is in fact *directly* replaced on the pre-mRNA prior to tri-snRNP addition, Prp45p with its overall charged and basic character, might actually be the more likely candidate to do this than Syf3p. If Prp45p binds to the pre-mRNA close to the branchpoint (and thus close to the 3' splice site), this would give another alternative for the time point of the observed Prp22p-Prp45p interaction, namely during 3' splice site selection and cleavage. However, the fact that in the Prp45p two-hybrid screen, no interactions with known second step splicing factors was observed argues against a direct involvement of the protein in the second step.

To date, it is not known whether Prp22p is associated with the spliceosome prior to step 1. An important experiment that needs to be done is therefore to use Prp22p or an epitope tagged form of it to try to coprecipitate spliceosomes which are stalled prior to the first transesterification reaction in order to provide evidence for Prp22p presence in the spliceosome at this stage. However, if there is an interaction of Prp45p with Prp22p during spliceosome assembly, it would likely be non-essential for splicing, since Prp22p is not required prior to step 1 (Schwer and Gross, 1998). The requirement for Prp45p is not dependent on the branchpoint-3' splice site distance, in contrast to the requirement for Prp22p, which is dispensable for splicing of short, but required for splicing of introns with large branchpoint-3' splice site distances, as discussed earlier. Hypothetically, a situation could be imagined in which Prp45p is bound in the region of the branch sequence, where it acts as a general factor during 3' splice site recognition and cleavage. Prp22p, by binding Prp45p, could form a bridge between the branchpoint and the 3' splice site in introns, in which the 3' splice site is far away from the branchpoint, thereby supporting 3' splice site recognition.

Whatever the exact function of Prp45p, the strong and reciprocal two-hybrid interaction with Prp46p suggests that the two novel splicing factors act in a concerted manner during pre-mRNA splicing.

Another line of experiments that should be followed is to verify the two-hybrid interactions identified in this work by testing the interactions in *in vitro* binding assays. In particular the Prp22p-Prp45p and Prp22p-Syf3p interactions should be tested again, since in these cases no reciprocal two-hybrid interactions (at least not in the screens) were observed. Furthermore, these interactions are the first data which suggest a link of Prp22p with factors that are required already prior to the first step and are thus important to be verified. *In vitro* binding studies could also reveal which of the two-hybrid interactions are direct, and which might be bridged by a third (or even more) protein(s). In particular for the triangular interactions Syf3p-Prp45p-Prp22p-Syf3 and -Syf3-Prp45p-Prp46p-Syf3 it could be possible, that one of the proteins is actually "sandwiched" in between the others, so that two of them do not contact each other directly.

Several observations suggest that phosphorylation events might influence the action of Prp22p, Prp45 and Prp46p: i) two putative phosphorylation sites are present within the only 30 amino acid long Prp22p-interacting region of Prp45p. ii) Prp45p shares similarity in its c-terminal half with *src*-like tyrosine kinases and furthermore contains 3 tyrosine residues which are 100% conserved among its putative homologues. iii) hPrp45p migrates as three distinct spots in 2D gels (Neubauer *et al.*, 1998), which could represent different phosphorylation states. iv) the *Arabidopsis* homologue of Prp46p, PRL1, interacts *in vitro* with the protein kinase C-beta II isoform (PKC-βII) (Nemeth *et al.*, 1998). v) both Prp46p and Prp45p interact with protein kinase C protein (Pkc1p; Ybl105p) in the two-hybrid screens (this work, A. Clark and C. Russell, unpublished). Therefore, initial studies should be performed to test a) whether Prp45p and Prp46p can be found phosphorylated in cell extracts and b) whether the proteins can be phosphorylated *in vitro*. If the proteins are indeed subjected to phosphorylation, deletion experiments or site-directed mutagenesis targeting the diverse putative phosphorylation sites should allow the determination of the residues involved. Thereafter, the effects of mutations within these residues on the interactions between the proteins and on pre-mRNA splicing could be investigated.

To hypothesise phosphorylation events is particularly intriguing with respect to the putative cell cycle link of the described network of proteins. The absence of functional Syf1p, Syf3p, Cef1p and Prp22p leads to a cell cycle arrest in G2-phase. It would be interesting to investigate the effects of Prp45p and Prp46p depletion on cell cycle progression, using the conditional strains constructed in this work. Depletion of prp5, the putative Prp46p-homologue in fission yeast, also causes the cells to arrest in late G2-phase (McDonald *et al.*, 1999). Therefore, it is likely that *S. cerevisiae* cells would face the same fate upon Prp46p or Prp45p depletion. Interestingly, the Pkc1 protein, identified in the Prp45p and Prp46p two-hybrid screens, is also required to promote progression of the cells from G2 to mitosis (Levin *et al.*, 1990). An investigation of whether indeed a splicing checkpoint exists, through which the cell cycle regulates gene expression or whether the failure to remove specific introns elicits a cell cycle checkpoint response that arrests the cell cycle could be a challenging long-term project.

References

Abovich, N., Liao, X. C. and Rosbash, M. The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes Dev* **8**, 843-54 (1994).

Abovich, N. and Rosbash, M. Cross-intron bridging interactions in the yeast commitment complex are conserved in mammals. *Cell* **89**, 403-12 (1997).

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403-10 (1990).

Anderson, G. J., Bach, M., Luehrmann, R. and Beggs, J. D. Conservation between yeast and man of a protein associated with U5 small nuclear ribonucleoprotein. *Nature* **342**, 819-21 (1989).

Ansari, A. and Schwer, B. SLU7 and a novel activity, SSF1, act during the PRP16-dependent step of yeast pre-mRNA splicing. *Embo Journal* **14**, 4001-9 (1995).

Anthony, J. G., Weidenhammer, E. M and Woolford, J. L. Jr. The yeast Prp3 protein is a U4/U6 snRNP protein necessary for integrity of the U4/U6 snRNP and the U4/U6.U5 tri-snRNP. *RNA* **3**, 1143-52 (1997).

Arst, H. N. Jr., Bignell, E. and Tilburn, J. Two new genes involved in signalling ambient pH in *Aspergillus nidulans*. *Mol Gen Genet* **245**, 787-90 (1994).

Arenas, J. E. and Abelson, J. N. The *Saccharomyces cerevisiae PRP21* gene product is an integral component of the prespliceosome. *Proc Natl Acad Sci U S A* **90**, 6771-5 (1993).

Ast, G. and Weiner, A. M. A novel U1/U5 interaction indicates proximity between U1 and U5 snRNAs during an early step of mRNA splicing. *RNA* **3**, 371-81 (1997).

Ayadi, L., Callebaut, I., Saguez, C., Villa ,T., Mornon, J. P. and Banroques, J. Functional and structural characterization of the prp3 binding domain of the yeast prp4 splicing factor. *J Mol Biol* **284**, 673-87 (1998).

Ayadi, L., Miller, M. and Banroques, J. Mutations within the yeast U4/U6 snRNP protein Prp4 affect a late stage of spliceosome assembly. *RNA* **3**, 197-209 (1997).

Banroques, J. and Abelson, J. N. PRP4: a protein of the yeast U4/U6 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **9**, 3710-9 (1989).

Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F. and Cullin, C. A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **21**, 3329-30 (1993).

Baudino, T. A., Kraichely, D. M., Jefcoat, S. C. Jr, Winchester, S. K., Partridge, N. C. and MacDonald, P.N. Isolation and characterization of a novel coactivator protein, NCoA-62, involved in vitamin D-mediated transcription. *J Biol Chem* **26**, 16434-41 (1998).

Beggs, J. D. Yeast splicing factors and genetic strategies for their analysis. In *Pre-mRNA Processing*, Lamond, A. I. (ed.), R. G. Landes Company, 79-95 (1995).

Beggs, J. D., Teigelkamp, S. and Newman, A. J. The role of PRP8 protein in nuclear pre-mRNA splicing in yeast. *J Cell Sci Suppl* **19**,101-5 (1995).

Behrens, S. E., Galisson, F., Legrain, P. and Luehrmann, R. Evidence that the 60-kDa protein of 17S U2 small nuclear ribonucleoprotein is immunologically and functionally related to the yeast PRP9 splicing factor and is required for the efficient formation of prespliceosomes. *Proc Natl Acad Sci U S A* **90**, 8229-33 (1993).

Bennett, M. and Reed, R. Correspondence between a mammalian spliceosome component and an essential yeast splicing factor. *Science* **262**, 105-8 (1993).

Ben-Yehuda, S., Dix, I., Russell, C. S., Levy, S., Beggs, J. D. and Kupiec, M. Identification and functional analysis of hPRP17, the human homologue of the *PRP17/CDC40* yeast gene involved in splicing and cell cycle control. *RNA* **4**, 1304-12 (1998).

Berglund, J. A., Chua, K., Abovich, N., Reed, R. and Rosbash, M. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **89**, 781-7 (1997).

Blanton, S., Srinivasan, A. and Rymond, B. C. PRP38 encodes a yeast protein required for pre-mRNA splicing and maintenance of stable U6 small nuclear RNA levels. *Mol Cell Biol* **12**, 3939-47 (1992).

Boeck, R., Lapeyre, B., Brown, C. E., Sachs, A. B. Capped mRNA degradation intermediates accumulate in the yeast spb8-2 mutant. *Mol Cell Biol* **18**, 5062-72 (1998).

Bordonne, R. and Tarassov, I. The yeast SME1 gene encodes the homologue of the human E core protein. *Gene* **176**, 111-7 (1996).

Branlant, C., Krol, A., Ebel, J. P., Lazar, E., Haendler, B. and Jacob, M. U2 RNA shares a structural domain with U1, U4, and U5 RNAs. *Embo Journal* **1**, 1259-65 (1982).

Brent, R. and Finley, R. L. Jr. Understanding gene and allele function with two-hybrid methods. *Annu Rev Genet* **31**, 663-704 (1997).

Brown, J. D. and Beggs, J. D. Roles of PRP8 protein in the assembly of splicing complexes. *Embo Journal* **11**, 3721-9 (1992).

Brown, S. J., Kellett, P. J. and Lippard, S. J. Ixr1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin. *Science* **261**, 603-5 (1993).

Brys, A. and Schwer, B. Requirement for SLU7 in yeast pre-mRNA splicing is dictated by the distance between the branchpoint and the 3' splice site. *RNA* **2**, 707-17 (1996).

Burge, C. B., Tuschl, T. and Sharp, P. A. Splicing of precursors to mRNAs by the spliceosome. In *The RNA World*; Gesteland, R. F., Cech, T. R and Atkins, J. F. (eds), Cold Spring Harbor Laboratory Press, 525-560 (1999).

Burgess, S., Couto, J. R. and Guthrie, C. A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell* **60**, 705-17 (1990).

Burgess, S. M., Guthrie, C. A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell* **73**, 1377-91 (1993).

Bycroft, M., Hubbard, T. J., Proctor, M., Freund, S. M. and Murzin, A. G. The solution structure of the S1 RNA binding domain: a member of an ancient nucleic acid-binding fold. *Cell* **88**, 235-42 (1997).

Caspary, F., Shevchenko, A., Wilm, M. and Seraphin, B. Partial purification of the yeast U2 snRNP reveals a novel yeast pre-mRNA splicing factor required for pre-spliceosome assembly. *Embo Journal* **18**, 3463-3474 (1999).

Chang, B.Y., Conroy, K. B., Machleder, E. M. and Cartwright, C. A. RACK1, a receptor for activated C kinase and a homolog of the beta subunit of G proteins, inhibits activity of src tyrosine kinases and growth of NIH 3T3 cells. *Mol Cell Biol* **18**, 3245-56 (1998).

Chen, H. -R., Jan, S. -P., Tsao, T. Y., Sheu, Y. -J., Banroques, J. and Cheng S. -C. Snt309p, a component of the Prp19p-associated complex that interacts with Prp19p and associates with the spliceosome simultaneously with or immediately after dissociation of U4 in the same manner as Prp19p. *Mol Cell Biol* **18**, 2196-204 (1998).

Chen, W. and Struhl, K. Yeast mRNA initiation sites are determined primarily by specific sequences, not by the distance from the TATA element. *Embo Journal* **4**, 3273-80 (1985).

Cherry, J. M., Ball, C., Weng, S., Juvik, G., Schmidt, R., Adler, C., Dunn, B., Dwight, S., Riles, L., Mortimer, R. K. and Botstein, D. Genetic and physical maps of *Saccharomyces cerevisiae*. *Nature* **387**(6632 Suppl) 67-73 (1997).

Chien, C. T., Bartel, P. L., Sternglanz, R. and Fields, S. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc Natl Acad Sci U S A* **88**, 9578-82 (1991).

Chua, K. and Reed, R. Human step II splicing factor hSlu7 functions in restructuring the spliceosome between the catalytic steps of splicing. *Genes Dev* **13**, 841-50 (1999).

Chung S., McLean M. R. and Rymond, B. C. Yeast ortholog of the *Drosophila* crooked neck protein promotes spliceosome assembly through stable U4/U6.U5 snRNP addition. *RNA* **5**, 1042-54 (1999).

Church, G. M. and Gilbert, W. Genomic sequencing. *Proc Natl Acad Sci U S A* **81**, 1991-5 (1984).

Company, M., Arenas, J. and Abelson, J. Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature* **349**, 487-93 (1991).

Cooper, M., Johnston, L. H. and Beggs, J. D. Identification and characterization of Uss1p (Sdb23p): a novel U6 snRNA-associated protein with significant similarity to core proteins of small nuclear ribonucleoproteins. *Embo Journal* **14**, 2066-75 (1995).

Courtneidge, S.A., Fumagalli, S., Koegl, M., Superti-Furga, G. and Twamley-Stein, G.M. The Src family of protein tyrosine kinases: regulation and functions. *Dev (Suppl)*, 57-64 (1993).

Das, A. K., Cohen, P. T. W. and Barford, D. The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *Embo Journal* **17**, 1192-99 (1998).

Daugeron, M. C. and Linder, P. Dbp7p, a putative ATP-dependent RNA helicase from *Saccharomyces cerevisiae*, is required for 60S ribosomal subunit assembly. *RNA* **4**, 566-81 (1998).

de la Cruz, J., Kressler, D. and Linder P. Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families. *TIBS* **24**, 192-98 (1999).

del Olmo, M., Mizrahi, N., Gross, S. and Moore, C. L. The Uba2 and Ufd1 proteins of *Saccharomyces cerevisiae* interact with poly(A) polymerase and affect the polyadenylation activity of cell extracts. *Mol Gen Genet* **255**, 209-18 (1997).

Devereux, J., Haeberli, P. and Smithies, O. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* **12**, 387-95 (1984).

Denison, S., Negrete-Urtasun, S., Mingot, J. M., Tilburn, J., Mayer, W. A., Goel, A., Espeso, E. A., Penalva, M. A. and Arst, H. N. Jr. Putative membrane components of signal transduction pathways for ambient pH regulation in *Aspergillus* and meiosis in *Saccharomyces* are homologous. *Mol Microbiol* **30**, 259-64 (1998).

Diehl, B.E. and Pringle, J.R. Molecular analysis of *Saccharomyces cerevisiae* chromosome I: identification of additional transcribed regions and demonstration that some encode essential functions. *Genetics* **127**, 287-98 (1991).

Dix, I., Russell, C., Yehuda, S. B., Kupiec, M. and Beggs, J. D. The identification and characterization of a novel splicing protein, Isy1p, of *Saccharomyces cerevisiae*. *RNA* **5**, 360-68 (1999).

Dix, I., Russell, C. S., O'Keefe, R. T., Newman, A. J. and Beggs, J. D. Protein-RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae*. *RNA* **4**, 1675-86 (1998).

Farruggio, D. C., Townsley, F. M. and Ruderman, J. V. Cdc20 associates with the kinase aurora2/Aik. *Proc Natl Acad Sci U S A* **96**, 7306-11 (1999).

Feinberg, A. P. and Vogelstein, B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**, 6-13 (1983).

Feinberg, A. P. and Vogelstein, B. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Addendum. *Anal Biochem* **137**, 266-7 (1984).

Fields, S. and Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-6 (1989).

Fleckner, J., Zhang, M., Valcarcel, J. and Green, M. R. U2AF65 recruits a novel human DEAD box protein required for the U2 snRNP-branchpoint interaction. *Genes Dev* **11**, 1864-72 (1997).

Folk, P., Puta, F., Krpejsova, L., Blahuskova, A., Markos, A., Rabino, M. and Dottin, R. P. The homolog of chromatin binding protein Bx42 identified in *Dictyostelium*. *Gene* **181**, 229-31 (1996).

Fortes, P., Kufel, J., Fornerod, M., Polycarpou-Schwarz, M., Lafontaine, D., Tollervey, D., Mattaj, I. W. Genetic and Physical Interactions Involving the Yeast Nuclear Cap-Binding Complex. *Mol Cell Biol* **19**, 6543-53 (1999).

Frank, D. and Guthrie, C. An essential splicing factor, *SLU7*, mediates 3' splice site choice in yeast. *Genes Dev* **6**, 2112-24 (1992).

Frank, D., Patterson, B. and Guthrie, C. Synthetic lethal mutations suggest interactions between U5 small nuclear RNA and four proteins required for the second step of splicing. *Mol Cell Biol* **12**, 5197-205 (1992).

Frasch, M. and Saumweber, H. Two proteins from *Drosophila* nuclei are bound to chromatin and are detected in a series of puffs on polytene chromosomes. *Chromosoma* **97**, 272-81(1989).

Fromont-Racine, M., Rain, J. C. and Legrain, P. Toward a functional analysis of the yeast genome through exhaustive two-hybrid screens. *Nat Genet* **16**, 277-82 (1997).

Futai, E., Maeda, T., Sorimachi, H., Kitamoto, K., Ishiura, S. and Suzuki, K. The protease activity of a calpain-like cysteine protease in *Saccharomyces cerevisiae* is required for alkaline adaptation and sporulation. *Mol Gen Genet* **260**, 559-68 (1999).

Galisson, F. and Legrain, P. The biochemical defects of prp4-1 and prp6-1 yeast splicing mutants reveal that the PRP6 protein is required for the accumulation of the [U4/U6.U5] tri-snRNP. *Nucleic Acids Res* **11**, 1555-62 (1993).

Gangloff, S., McDonald, J. P., Bendixen, C., Arthur, L. and Rothstein, R. The yeast type I topoisomerase Top3 interacts with Sgs1, a DNA helicase homolog: a potential eukaryotic reverse gyrase. *Mol Cell Biol* **14**, 8391-98 (1994).

Gietz, D., St. Jean, A., Woods, R. A. and Schiestl, R. H. Improved method for high efficiency transformation of intact yeast cells. *Nucleic Acids Res* **20**, 1425 (1992).

Goebl, M. and Yanagida, M. The TPR snap helix: a novel protein repeat motif from mitosis to transcription.*TIBS* **16**, 173-7 (1991).

Gondran, P. and Dautry, F. Regulation of mRNA splicing and transport by the tyrosine kinase activity of src. *Oncogene* **18**, 2547-55 (1999).

Gottschalk, A., Tang, J., Puig, O., Salgado, J., Neubauer, G., Colot, H. V., Mann, M., Seraphin, B., Rosbash, M., Luehrmann, R. and Fabrizio, P. A comprehensive biochemical and genetic analysis of the yeast U1 snRNP reveals five novel proteins. *RNA* **4**, 374-93 (1998).

Graack, H. R. and Wittmann-Liebold, B. Mitochondrial ribosomal proteins (MRPs) of yeast. *Biochem J* **329**, 433-48 (1998).

Grauslund, M., Didion, T., Kielland-Brandt, M. C. and Andersen, H. A. *BAP2*, a gene encoding a permease for branched-chain amino acids in *Saccharomyces cerevisiae*. *Biochim Biophys Acta* **1269**, 275-80 (1995).

Gribskov, M. Translational initiation factors IF-1 and eIF-2 alpha share an RNA-binding motif with prokaryotic ribosomal protein S1 and polynucleotide phosphorylase. *Gene* **119**, 107-11 (1992).

Guthrie, C. Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science* **253**, 157-63 (1991).

Guthrie, C. and Patterson, B. Spliceosomal snRNAs. *Annu Rev Genet* **22**, 387-419 (1988).

Hamm, J. and Lamond, A. I. Spliceosome assembly: the unwinding role of DEAD-box proteins. *Curr Biol* **8**, 532-4 (1998).

Hamm, J., van Santen, V. L., Spritz, R. A. and Mattaj, I. W. Loop I of U1 small nuclear RNA is the only essential RNA sequence for binding of specific U1 small nuclear ribonucleoprotein particle proteins. *Mol Cell Biol* **8**, 4787-91 (1988).

Horowitz, D. S. and Abelson, J. A U5 small nuclear ribonucleoprotein particle protein involved only in the second step of pre-mRNA splicing in *Saccharomyces cerevisiae*. *Mol Cell Biol* **13**, 2959-70 (1993).

Horowitz, D. S. and Krainer, A. R. A human protein required for the second step of pre-mRNA splicing is functionally related to a yeast splicing factor. *Genes Dev* **11**, 139-51 (1997).

Hermann, H., Fabrizio, P., Raker, V. A., Foulaki, K., Hornig, H., Brahms, H. and Hoey, T., Weinzierl. R. O., Gill, G., Chen, J.L., Dynlacht, B.D. and Tjian, R. Molecular cloning and functional analysis of Drosophila TAF110 reveal properties expected of coactivators. *Cell* **29,** 247-60 (1993).

Hollenberg, S. M., Sternglanz, R., Cheng, P. F., Weintraub, H. Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system. *Mol Cell Biol* **15**, 3813-22 (1995).

Hwang, L. H. and Murray, A. W. A novel yeast screen for mitotic arrest mutants identifies DOC1, a new gene involved in cyclin proteolysis. *Mol Biol Cell* **8**, 1877-87 (1997).

Icho, T. and Wickner, R. B. The MAK11 protein is essential for cell growth and replication of M double-stranded RNA and is apparently a membrane-associated protein. *J Biol Chem* **263**, 1467-75 (1988).

Jacquier, A. Self-splicing group II and nuclear pre-mRNA introns: how similar are they? *TIBS* **15**, 351-4 (1990).

Jarmolowski, A. and Mattaj, I. W. The determinants for Sm protein binding to Xenopus U1 and U5 snRNAs are complex and non-identical. *Embo Journal* **12**, 223-32 (1993).

Jones, M. H., Frank, D. N. and Guthrie, C. Characterization and functional ordering of Slu7p and Prp17p during the second step of pre-mRNA splicing in yeast. *Proc Natl Acad Sci U S A* **92**, 9687-91 (1995).

Jorgensen, M. U., Bruun, M. B., Didion, T. and Kielland-Brandt, M. C. Mutations in five loci affecting GAP1-independent uptake of neutral amino acids in yeast.*Yeast* **14**, 103-14 (1998).

Katz, R. A. and Jentoft, J. E. What is the role of the cys-his motif in retroviral nucleocapsid (NC) proteins? *Bioessays* **11**, 176-81 (1989).

Kambach, C. and Mattaj, I. W. Intracellular distribution of the U1A protein depends on active transport and nuclear binding to U1 snRNA. *J Cell Biol* **118**, 11-21 (1992).

Kao, H. Y. and Siliciano, P. G. The yeast homolog of the U1 snRNP protein 70K is encoded by the SNP1 gene. *Nucleic Acids Res* **20**, 4009-13 (1992).

Kao, H. Y. and Siliciano, P. G. Identification of Prp40, a novel essential yeast splicing factor associated with the U1 small nuclear ribonucleoprotein particle. *Mol Cell Biol* **16**, 960-7 (1996).

Kessler, M. M., Henry, M. F., Shen, E., Zhao, J., Gross, S., Silver, P. A. and Moore, C. L. Hrp1, a sequence-specific RNA-binding protein that shuttles between the nucleus and the cytoplasm, is required for mRNA 3'-end formation in yeast. *Genes Dev* **11**, 2545-56 (1997).

Kim, S. H. and Lin, R. J. Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. *Proc Natl Acad Sci U S A* **90**, 888-92 (1993).

Kim, D. H. and Rossi, J. J. The first ATPase domain of the yeast 246-kDa protein is required for *in vivo* unwinding of the U4/U6 duplex. *RNA* **5**, 959-71 (1999).

King, D. S. and Beggs, J. D. Interactions of PRP2 protein with pre-mRNA splicing complexes in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **18**, 6559-64 (1990).

Kraemer, A. The biochemistry of pre-mRNA splicing. In *Pre-mRNA Processing*, Lamond, A. I. (ed.), R. G. Landes Company, 35-64 (1995).

Kraemer, A., Legrain, P., Mulhauser, F., Groning, K., Brosi, R. and Bilbe, G. Splicing factor SF3a60 is the mammalian homologue of PRP9 of *S.cerevisiae*: the conserved zinc finger-like motif is functionally exchangeable *in vivo*. *Nucleic Acids Res* **22**, 5223-8 (1994).

Kraemer, A., Mulhauser, F., Wersig, C., Groning, K. and Bilbe, G. Mammalian splicing factor SF3a120 represents a new member of the SURP family of proteins and is homologous to the essential splicing factor PRP21p of *Saccharomyces cerevisiae*. *RNA* **3**, 260-72 (1995).

Krainer, A. R., Maniatis, T., Ruskin, B. and Green, M. R. Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced *in vitro*. *Cell* **36**, 993-1005 (1984).

Kuhn, A. N., Li, Z. and Brow, D. A. Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. *Mol Cell* **3**, 65-75 (1999).

Kyte, J. and Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32 (1982).

Laggerbauer, B., Achsel, T. and Luehrmann, R. The human U5-200kD DEXH-box protein unwinds U4/U6 RNA duplices *in vitro*. *Proc Natl Acad Sci U S A* **95**, 4188-92 (1998).

Lafontaine, D. and Tollervey, D. One-step PCR mediated strategy for the construction of conditionally expressed and epitope tagged yeast proteins. *Nucleic Acids Res* **24**, 3469-72 (1996).

Lamb, J. R., Tugendreich, S. and Hieter, P. Tetratrico peptide repeat interactions: to TPR or not to TPR? *TIBS* **20**, 257-59 (1995).

Larsen, J., Beug, H. and Hayman, M. J. The v-ski oncogene cooperates with the v-sea oncogene in erythroid transformation by blocking erythroid differentiation. *Oncogene* **7,** 1903-11 (1992).

Lauber, J., Plessel, G., Prehn, S., Will, C. L., Fabrizio, P., Groning, K., Lane, W. S. and Luehrmann, R. The human U4/U6 snRNP contains 60 and 90kD proteins that are structurally homologous to the yeast splicing factors Prp4p and Prp3p. *RNA* **3**, 926-41 (1997).

Legrain, P. and Chapon, C. Interaction between PRP11 and SPP91 yeast splicing factors and characterization of a PRP9-PRP11-SPP91 complex. *Science* **262**, 108-10 (1993).

Legrain, P. and Choulika, A. The molecular characterization of PRP6 and PRP9 yeast genes reveals a new cysteine/histidine motif common to several splicing factors. *Embo Journal* **9**, 2775-81 (1990).

Legrain, P., Seraphin, B. and Rosbash, M. Early commitment of yeast pre-mRNA to the spliceosome pathway. *Mol Cell Biol* **8**, 3755-60 (1988).

Lerner, M. R. and Steitz, J. A. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc Natl Acad Sci U S A* **76**, 5495-9 (1979).

Levin, D. E., Fields, F. O., Kunisawa, R., Bishop, J. M. and Thorner, J. A candidate protein kinase C gene, PKC1, is required for the *S. cerevisiae* cell cycle. *Cell* **62**, 213-24 (1990).

Li, W. and Mitchell, A. P. Proteolytic activation of Rim1p, a positive regulator of yeast sporulation and invasive growth. *Genetics* **145**, 63-73 (1997).

Li, Z. and Brow, D. A. A spontaneous duplication in U6 spliceosomal RNA uncouples the early and late functions of the ACAGA element in vivo. *RNA* **2**, 879-94 (1996).

Lin, R. J., Newman, A. J., Cheng, S. -C. and Abelson, J. Yeast m-RNA splicing *in vitro*. *J Biol Chem* **260**, 14780-92 (1985).

Luehrmann, R. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. *Embo Journal* **14**, 2076-88 (1995).

Luukkonen, B. G. and Seraphin, B.
The role of branchpoint-3' splice site spacing and interaction between intron terminal nucleotides in 3' splice site selection in *Saccharomyces cerevisiae*. *Embo Journal* **16**, 779-92 (1997).

Lundgren, K., Allan, S., Urushiyama, S., Tani, T., Ohshima, Y., Frendewey, D. and Beach, D. A connection between pre-mRNA splicing and the cell cycle in fission yeast: *cdc28+* is allelic with *prp8+* and encodes an RNA-dependent ATPase/helicase. *Mol Biol Cell* **7**, 1083-94 (1996).

Lustig, A. J., Lin, R. J. and Abelson, J. The yeast RNA gene products are essential for mRNA splicing *in vitro*. *Cell* **47**, 953-63 (1986).

Maddock, J. R., Roy, J. and Woolford, Jr. J. L. Six novel genes necessary for pre-mRNA splicing in *Saccharomyces cerevisiae*. *Nucleic Acids Res* **24**, 1037-44, (1996).

Mann, C., Micouin, J. Y., Chiannilkulchai, N., Treich, I., Buhler, J. M. and Sentenac, A. *RPC53* encodes a subunit of *Saccharomyces cerevisiae* RNA polymerase C (III) whose inactivation leads to a predominantly G1 arrest. *Mol Cell Biol* **12**, 4314-26 (1992).

Maschhoff, K. L., Padgett, R. A. The stereochemical course of the first step of pre-mRNA splicing. *Nucleic Acids Res* **21**, 5456-62 (1993).

Mayes, A. E. Identification and characterisation of proteins interacting with Uss1p: A *Saccharomyces cerevisiae* Sm-like protein. PhD Thesis, University of Edinburgh (1999).

Mayes, A. E., Verdone L., Legrain P. and Beggs, J. D. Characterisation of Sm-like proteins in yeast and their association with U6 snRNA. *Embo Journal* **18**, 4321-31 (1999).

McA'Nulty, M. M., Whitehead, J. P. and Lippard, S. J. Binding of Ixr1, a yeast HMG-domain protein, to cisplatin-DNA adducts *in vitro* and *in vivo*. *Biochemistry* **35**, 6089-99 (1996).

McDonald, W. H., Ohi, R., Smelkova, N., Frendewey, D. and Gould, K. L. Myb-related fission yeast cdc5p is a component of a 40S snRNP-containing complex and is essential for pre-mRNA splicing. *Mol Cell Biol* **19**, 5352-62 (1999).

Michel, F. and Ferat, J. L. Structure and activities of group II introns. *Annu Rev Biochem* **64**, 435-61 (1995).

Minvielle-Sebastia, L., Beyer, K., Krecic, A. M., Hector, R. E., Swanson, M. S. and Keller, W. Control of cleavage site selection during mRNA 3'-end formation by a yeast hnRNP. *Embo Journal* **17**, 7454-68 (1998).

Minvielle-Sebastia, L., Preker, P. J. and Keller, W. RNA14 and RNA15 proteins as components of a yeast pre-mRNA 3'-end processing factor. *Science* **266**, 1702-5 (1994).

Mongiovi, A.M., Romano, P. R., Panni, S., Mendoza, M., Wong, W. T., Musacchio, A., Cesareni, G. and Di Fiore, P. A novel peptide-SH3 interaction. *Embo Journal* **18**, 5300-9 (1999).

Moore, M. J.and Sharp, P. A. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature* **365**, 364-8 (1993).

Noble, S. M. and Guthrie, C. Identification of novel genes required for yeast pre-mRNA splicing by means of cold-sensitive mutations. *Genetics* **143**, 67-80 (1996).

Natorff, R., Piotrowska, M. and Paszewski, A. The *Aspergillus* nidulans sulphur regulatory gene sconB encodes a protein with WD40 repeats and an F-box. *Mol Gen Genet* **257**, 255-63 (1998).

Neel, H., Gondran, P., Weil, D.and Dautry, F. Regulation of pre-mRNA processing by src. *Curr Biol* **5,** 413-22 (1995).

Negrete-Urtasun, S., Denison, S. and Arst, H. N. Jr. Characterisation of the pH signal transduction pathway gene palA of *Aspergillus nidulans* and identification of possible homologs. *J Bacteriol* **179**, 1832-35 (1997).

Nelissen, R. L., Will, C. L., van Venrooij, W. J. and Luehrmann, R. The association of the U1-specific 70K and C proteins with U1 snRNPs is mediated in part by common U snRNP proteins. *Embo Journal* **13**, 4113-25 (1994).

Nemeth, K., Salchert, K., Putnoky, P., Bhalerao, R., Koncz-Kalman, Z., Stankovic-Stangeland, B., Bako, L., Mathur, J., Okresz, L., Stabel, S., Geigenberger, P., Stitt, M., Redei, G. P., Schell, J. and Koncz, C. Pleiotropic control of glucose and hormone responses by PRL1, a nuclear WD protein, in *Arabidopsis*. *Genes Dev* **12**, 3059-73 (1998).

Neubauer, G., Gottschalk, A., Fabrizio, P., Seraphin, B., Luehrmann, R. and Mann, M. Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proc Natl Acad Sci U S A* **94**, 385-90 (1997).

Neubauer, G., King, A., Rappsilber, J., Calvio, C., Watson, M., Ajuh, P., Sleeman, J., Lamond, A. and Mann, M. Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat Genet* **20**, 46-50 (1998).

Newman, A. J. and Norman, C. Mutations in the U5 snRNA alter the specificity of 5' splice site cleavage. *Cell* **65**, 115-123 (1991).

Newman, A. J. and Norman, C. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell* **68**, 743-54 (1992).

Newman, A. J. Small nuclear RNAs and pre-mRNA splicing. *Curr Opin Cell Biol* **6**, 360-367 (1994).

Newman, A. J., Teigelkamp, S. and Beggs, J. D. snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. *RNA* **1**, 968-80 (1995).

Ng, S. W., Liu, Y., Hasselblatt, K. T., Mok, S. C. and Berkowitz, R. S. A new human topoisomerase III that interacts with SGS1 protein. *Nucl Acid Res* **27**, 993-1000 (1999).

Nilsen, T. W. RNA-RNA interactions in nuclear pre-mRNA splicing. In Simons, R. W. and Grunberg-Manago, M. (eds), *RNA structure and function*. Cold Spring Harbour Laboratory Press, 279-307 (1998).

O'Day, C. L., Dalbadie-McFarland, G. and Abelson, J. The *Saccharomyces cerevisiae* Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. *J Biol Chem* **271**, 33261-7 (1996).

Ohi, R., Feoktistova, A., McCann, S., Valentine, V., Look, A. T., Lipsick, J. S. and Gould, K. L. Myb-related *Schizosaccharomyces pombe* cdc5p is structurally and functionally conserved in eukaryotes. *Mol Cell Biol* **18**, 4097-108 (1998).

Ohno, M. and Shimura, Y. A human RNA helicase-like protein, HRH1, facilitates nuclear export of spliced mRNA by releasing the RNA from the spliceosome. *Genes Dev* **10**, 997-1007 (1996).

O'Keefe, R. T., Norman, C. and Newman, A. J. The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. *Cell* **86**, 679-89 (1996).

Orejas, M., Espeso, E. A., Tilburn, J., Sarkar, S., Arst, H. N. Jr. and Penalva, M. A. Activation of the *Aspergillus* PacC transcription factor in response to alkaline ambient pH requires proteolysis of the carboxy-terminal moiety. *Genes Dev* **9**, 1622-32 (1995).

Padgett, R. A., Konarska, M. M., Grabowski, P. J., Hardy, S. F and Sharp P. A. Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. *Science* **225**, 898-903 (1984).

Parker, R., Siliciano, P. G. and Guthrie, C. Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell* **49**, 229-39 (1987).

Parker, R. and Siliciano, P. G. Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. *Nature* **361**, 660-2 (1993).

Patterson, B. and Guthrie, C. A U-rich tract enhances usage of an alternative 3' splice site in yeast. *Cell* **64**, 181-7 (1991).

Pause, A. and Sonenberg, N. Mutational analysis of a DEAD box RNA helicase: the mammalian translation initiation factor eIF-4A. *Embo Journal* **11**, 2643-54 (1992).

Pearson, W. R. and Lipman, D. J. Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* **85**, 2444-8 (1988).

Pikielny C. W. and Rosbash, M. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell* **41**, 119-26 (1985).

Plessel, G., Fischer, U. and Luehrmann, R. m3G cap hypermethylation of U1 small nuclear ribonucleoprotein (snRNP) in vitro: evidence that the U1 small nuclear RNA-(guanosine-N2)-methyltransferase is a non-snRNP cytoplasmic protein that requires a binding site on the Sm core domain. *Mol Cell Biol* **14**, 4160-72 (1994).

Plumpton, M., McGarvey, M. and Beggs, J. D. A dominant negative mutation in the conserved RNA helicase motif 'SAT' causes splicing factor PRP2 to stall in spliceosomes. *Embo Journal* **13**, 879-87 (1994).

Potashkin, J., Kim, D., Fons, M., Humphrey, T. and Frendeway, D. Cell-division-cycle defects associated with fission yeast pre-mRNA splicing mutants. *Curr Genet* **34**, 153-63 (1998).

Puig, O., Gottschalk, A., Fabrizio, P. and Seraphin, B. Interaction of the U1 snRNP with nonconserved intronic sequences affects 5' splice site selection. *Genes Dev* **13**, 569-80 (1999).

Pryer, N. K., Salama, N. R., Schekman, R. and Kaiser, C. A. Cytosolic Sec13p complex is required for vesicle formation from the endoplasmic reticulum *in vitro*. *J Cell Biol* **120**, 865-75 (1993).

Query, C. C., Moore, M. J. and Sharp, P. A. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev* **8**, 587-97 (1994).

Raker, V. A., Plessel, G. and Luehrmann, R. The snRNP core assembly pathway: identification of stable core protein heteromeric complexes and an snRNP subcore particle *in vitro*. *Embo Journal* **15**, 2256-69 (1996).

Rain, J. C., Rafi, Z., Rhani, Z., Legrain, P. and Kraemer, A. Conservation of functional domains involved in RNA binding and protein-protein interactions in human and *Saccharomyces cerevisiae* pre-mRNA splicing factor SF1. *RNA* **4**, 551-65 (1998).

Reddy, R. and Busch, H. Small RNAs: RNA sequences, structure and modifications. In *Structure and function of major and minor small ribonucleoprotein particles.* Birnstiel, M. L. (ed.), Springer Verlag Berlin, 1-37 (1988).

Romac, J. M., Graff, D. H and Keene, J. D. The U1 small nuclear ribonucleoprotein (snRNP) 70K protein is transported independently of U1 snRNP particles via a nuclear localization signal in the RNA-binding domain. *Mol Cell Biol* **14**, 4662-70 (1994).

Ron, D., Chen, C. H., Caldwell, J., Jamieson, L., Orr, E. and Mochly-Rosen, D. Cloning of an intracellular receptor for protein kinase C: a homolog of the beta subunit of G proteins. *Proc Natl Acad Sci U S A* **91**, 839-43 (1994).

Roy, J., Kim, K., Maddock, J. R., Anthony, J. G. and Woolford, J. L. Jr. The final stages of spliceosome maturation require Spp2p that can interact with the DEAH box protein Prp2p and promote step 1 of splicing. *RNA* **1**, 375-90 (1995).

Ruby, S. W., Chang, T. H. and Abelson, J. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes Dev* **7**, 1909-25 (1993).

Ruskin, B., Krainer, A. R., Maniatis, T. and Green, M. R. Excision of an intact intron as a novel lariat structure during pre-mRNA splicing *in vitro. Cell* **38**, 317-31 (1984).

Russnak, R., Nehrke, K. W. and Platt, T. *REF2* encodes an RNA-binding protein directly involved in yeast mRNA 3'- end formation. *Mol Cell Biol* **15**, 1689-97 (1995).

Russnak, R., Pereira, S. and Platt, T. RNA binding analysis of yeast REF2 and its two-hybrid interaction with a new gene product, FIR1. *Gene Expr* **6**, 241-58 (1996).

Rutz, B. and Seraphin, B. Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *RNA* **5**, 819-31 (1999).

Rymond, B. C. and Rosbash, M. Yeast pre-mRNA splicing. In *The molecular and cellular biology of the yeast Saccharomyces: gene expression.* Jones E. W., Pringle, J. R. and Broach, J. R. (eds); Cold Spring Harbor Laboratory Press, 143-92 (1992).

Rymond, B. C., Rokeach, L. A. and Hoch, S. O. Human snRNP polypeptide D1 promotes pre-mRNA splicing in yeast and defines nonessential yeast Smd1p sequences. *Nucleic Acids Res* **21**, 3501-5 (1993).

Salgado-Garrido, J., Bragado-Nilsson, E., Kandels-Lewis, S. and Seraphin, B. Sm and Sm-like proteins assemble in two related complexes of deep evolutionary origin. *Embo Journal* **18**, 3451-62 (1999).

Schappert, K. and Friesen, J. D. Genetic studies of the *PRP11* gene of *Saccharomyces cerevisiae. Mol Gen Genet* **226**, 277-82 (1991).

Schmitt, M. E., Brown, T. A. and Trumpower, B. L. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae. Nucleic Acids Res* **18**, 3091-92 (1990).

Schwer, B. and Gross, C. H. Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *Embo Journal* **17**, 2086-94 (1998).

Schwer, B. and Guthrie, C. PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. *Nature* **349**, 494-9 (1991).

Schwer, B. and Guthrie, C. A dominant negative mutation in a spliceosomal ATPase affects ATP hydrolysis but not binding to the spliceosome. *Mol Cell Biol* **12**, 3540-7 (1992).

Seraphin, B. Sm and Sm-like proteins belong to a large family: identification of proteins of the U6 as well as the U1, U2, U4 and U5 snRNPs. *Embo Journal* **14**, 2089-98 (1995).

Seraphin, B. and Rosbash, M. Mutational analysis of the interactions between U1 small nuclear RNA and pre-mRNA of yeast. *Gene* **82**, 145-51 (1989).

Seraphin, B. and Rosbash, M. The yeast branchpoint sequence is not required for the formation of a stable U1 snRNA-pre-mRNA complex and is recognized in the absence of U2 snRNA. *Embo Journal* **10**, 1209-16 (1991).

Shah, A. J., Tilburn, J., Adlard, M. W. and Arst, H. N. Jr. pH regulation of penicillin production in *Aspergillus nidulans*. *FEMS Microbiol Lett* **77**, 209-12 (1991).

Shea, J. E., Toyn, J. H. and Johnston, L. H. The budding yeast U5 snRNP Prp8 is a highly conserved protein which links RNA splicing with cell cycle progression. *Nucl Acid Res* **22**, 5555-64 (1994).

Smeekens, S. and Rook, F. Sugar sensing and sugar mediated signal transduction in plants. *Plant physiol* **115**, 7-13 (1997).

Smith, C. W., Porro, E. B., Patton, J. G. and Nadal-Ginard, B. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature* **342**, 243-7 (1989).

Smith, D. J., Proudfoot, A., Friedli, L., Klig, L. S., Paravicini, G. and Payton, M. A. PMI40, an intron-containing gene required for early steps in yeast mannosylation. *Mol Cell Biol* **12**, 2924-30 (1992).

Smith, P. A. Protein-protein interactions of Prp2p: a pre-mRNA splicing factor of *Saccharomyces cerevisiae*. PhD Thesis, University of Edinburgh (1999).

Smith, T. F., Gaitatzes, C., Saxena, K. and Neer, E. The WD-repeat: a common architecture for diverse functions. *TIBS* **24**, 181-85 (1999).

Smith, V. and Barrell, B. G. Cloning of a yeast U1 snRNP 70K protein homologue: functional conservation of an RNA-binding domain between humans and yeast. *Embo Journal* **10**, 2627-34 (1991).

Staley, J. P. and Guthrie, C. An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. *Mol Cell* **3**, 55-64 (1999).

Stevens, S. W. and Abelson, J. Purification of the yeast U4/U6.U5 small nuclear ribonucleoprotein particle and identification of its proteins. *Proc Natl Acad Sci U S A* **96**, 7226-31 (1999).

Struhl, K., Stinchcomb, D. T., Scherer, S. and Davis. R. W. High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. *Proc Natl Acad Sci U S A* **76**, 1035-9 (1979).

Su, S. S. and Mitchell, A. P. Identification of functionally related genes that stimulate early meiotic gene expression in yeast. *Genetics* **133**, 67-77 (1993).

Takagaki, Y. and Manley, J. L. A human polyadenylation factor is a G protein beta-subunit homologue. *J Biol Chem* **267**, 23471-4 (1992).

Tang, J., Abovich, N., Fleming, M. L., Seraphin, B. and Rosbash, M. Identification and characterization of a yeast homolog of U1 snRNP-specific protein C. *Embo Journal* **16**, 4082-91 (1997).

Tarn, W. -Y., Hsu, C. -H., Huang, K. -T., Chen, H. -R., Kao, H. -Y., Lee, K. -R. and Cheng, S. -C. Functional association of essential splicing factor(s) with PRP19 in a protein complex. *Embo Journal* **13**, 2421-31 (1994).

Tarn, W. -Y., Lee, K. -R. and Cheng, S. -C. The yeast PRP19 protein is not tightly associated with small nuclear RNAs, but appears to associate with the spliceosome after binding of U2 to the pre-mRNA and prior to formation of the functional spliceosome. *Mol Cell Biol* **13**, 1883-91 (1993a).

Tarn, W. -Y., Lee, K. -R. and Cheng, S. -C. Yeast precursor mRNA processing protein PRP19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. *Proc Natl Acad Sci U S A* **90**, 10821-5 (1993b).

Teigelkamp, S., McGarvey, M., Plumpton, M. and Beggs, J. D. The splicing factor PRP2, a putative RNA helicase, interacts directly with pre-mRNA. *Embo Journal* **13**, 888-97 (1994).

Teigelkamp, S., Newman, A. J. and Beggs, J. D. Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *Embo Journal* **14**, 2602-12 (1995).

Teigelkamp, S., Whittaker, E. and Beggs, J. D. Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. *Nucleic Acids Res* **23**, 320-26 (1995).

Thorsness, P. E., White, K. H. and Fox, T. D. Inactivation of YME1, a member of the ftsH-SEC18-PAS1-CDC48 family of putative ATPase-encoding genes, causes increased escape of DNA from mitochondria in *Saccharomyces cerevisiae*. *Mol Cell Biol* **13**, 5418-26 (1993).

Tilburn, J., Sarkar, S., Widdick, D. A., Espeso, E. A., Orejas, M., Mungroo, J., Penalva, M. A. and Arst, H. N. Jr. The *Aspergillus* PacC zinc finger transcription factor mediates regulation of both acid- and alkaline-expressed genes by ambient pH. *Embo Journal* **14**, 779-90 (1995).

Tomoyasu, T., Yuki, T., Morimura, S., Mori, H., Yamanaka, K., Niki, H., Hiraga, S. and Ogura, T. The *Escherichia coli* FtsH protein is a prokaryotic member of a protein family of putative ATPases involved in membrane functions, cell cycle control, and gene expression. *J Bacteriol* **175**, 1344-51 (1993).

Transy, C. and Legrain, P. The two-hybrid: an *in vivo* protein-protein interaction assay. *Mol Biol Rep* **21**, 119-27 (1995).

Tsai, W. Y., Chow, Y. T., Chen, H. R., Huang, K. T., Hong, R. I., Jan, S. P., Kuo, N. Y., Tsao, T. Y., Chen, C. H. and Cheng, S. C. Cef1p Is a component of the Prp19p-associated complex and essential for pre-mRNA splicing. *J Biol Chem* **274**, 907-13 (1999).

Umen, J. G. and Guthrie, C. Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* **1**, 584-97 (1995a).

Umen, J. G. and Guthrie, C. The second catalytic step of pre-mRNA splicing. *RNA* **1**, 869-85 (1995b).

Valcarcel, J., Singh, R. and Green, M. R. Mechanisms of regulated pre-mRNA splicing. In *pre-mRNA processing* edited by Lamond, A.I., R.G Landes Company (1995).

Vijayraghavan, U. and Abelson, J. PRP18, a protein required for the second reaction in pre-mRNA splicing. *Mol Cell Biol* **10**, 324-32 (1990).

Vijayraghavan, U., Company, M. and Abelson, J. Isolation and characterization of pre-mRNA splicing mutants of *Saccharomyces cerevisiae. Genes Dev* **3**, 1206-16 (1989).

Vijayraghavan, U., Parker, R., Tamm, J., Iimura, Y., Rossi, J., Abelson, J. and Guthrie, C. Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. *Embo Journal* **5**, 1683-95 (1986).

Wagner, J. D., Jankowsky, E., Company, M., Pyle, A. M.and Abelson, J.N. The DEAH-box protein PRP22 is an ATPase that mediates ATP-dependent mRNA release from the spliceosome and unwinds RNA duplexes. *Embo Journal* **15**, 2926-37 (1998).

Waksman, G., Kominos, D., Robertson, S.C., Pant, N., Baltimore, D., Birge, R.B., Cowburn, D., Hanafusa, H., Mayer, B.J., Overduin, M., Resh, M. D., Rios, C. B., Silverman, L. and Kuriyan J. Crystal structure of the phosphotyrosine recognition domain SH2 of v-src complexed with tyrosine-phosphorylated peptides. *Nature* **358**, 646-53 (1992).

Wall, M. A., Coleman, D. E., Lee, E., Iniguez-Lluhi, J. A., Posner, B. A., Gilman, A. G and Sprang, S. R.The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2. *Cell* **83**, 1047-58 (1995).

Wang, A., Forman-Kay, J., Luo, Y., Luo, M., Chow, Y. H., Plumb, J., Friesen, J. D., Tsui, L. C., Heng, H. H., Woolford, J. L. Jr. and Hu, J. Identification and characterization of human genes encoding Hprp3p and Hprp4p, interacting components of the spliceosome. *Hum Mol Genet* **6**, 2117-26 (1997).

Wang, Y., Wagner, J. D. and Guthrie, C. The DEAH-box splicing factor Prp16 unwinds RNA duplexes *in vitro. Curr Biol* **8**, 441-51 (1998).

Weidenhammer, E. M., Ruiz-Noriega, M. and Woolford, J. L. Jr. Prp31p promotes the association of the U4/U6 x U5 tri-snRNP with prespliceosomes to form spliceosomes in *Saccharomyces cerevisiae. Mol Cell Biol* **17**, 3580-8 (1997).

Wells, S. E. and Ares, M. Jr. Interactions between highly conserved U2 small nuclear RNA structures and Prp5p, Prp9p, Prp11p, and Prp21p proteins are required to ensure integrity of the U2 small nuclear ribonucleoprotein in *Saccharomyces cerevisiae. Mol Cell Biol* **14**, 6337-49 (1994).

Wieland, C., Mann, S., von Besser, H. and Saumweber, H. The *Drosophila* nuclear protein Bx42, which is found in many puffs on polytene chromosomes, is highly charged. *Chromosoma* **101**, 517-25 (1992).

Winkelmann, G., Bach, M. and Luehrmann, R. Evidence from complementation assays *in vitro* that U5 snRNP is required for both steps of mRNA splicing. *Embo Journal* **8**, 3105-12 (1989).

Xie, J., Beickman, K., Otte, E. and Rymond, B. C. Progression through the spliceosome cycle requires Prp38p function for U4/U6 snRNA dissociation. *Embo Journal* **17**, 2938-46 (1998).

Zhang, D. and Rosbash, M. Identification of eight proteins that cross-link to pre-mRNA in the yeast commitment complex. *Genes Dev* **13**, 581-92 (1999).

Zhang, X. and Schwer, B. Functional and physical interaction between the yeast splicing factors Slu7 and Prp18. *Nucleic Acids Res* **25**, 2146-52 (1997).

Zhou, C., Yang, Y. and Jong, A. Y. Mini-prep in ten minutes. *Biotechniques* **8**, 172-73 (1990).

Zhou, Z. and Reed, R. Human homologs of yeast Prp16 and Prp17 reveal conservation of the mechanism for catalytic step II of pre-mRNA splicing. *Embo Journal* **17**, 2095-106 (1998).

Zieve, G. W., Sauterer, R. A. Cell biology of the snRNP particles. *Crit Rev Biochem Mol Biol* **25**, 1-46 (1990).