Molecular Genetic Studies of the Prp8 Splicing Factor

Andrew James Brown

A Thesis Presented for the Degree of Doctor of Philosophy

University of Edinburgh

September 1995

Declaration

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

September 1995

Abstract

Introns which interrupt transcripts from eukaryotic genes are removed in a splicing reaction. Introns are recognised by particles called snRNPs, which assemble into a spliceosome complex in which the reactions of splicing occur. RNA components of the snRNPs are present at the spliceosome active site, and splicing is believed to be catalysed, at least in part, by RNA.

A factor which seems to have roles throughout spliceosome assembly and catalysis is the Prp8 protein of the budding yeast *Saccharomyces cerevisiae*. Prp8p is a component of the U5 snRNP important for entry of U5 and other snRNPs into the spliceosome, and is present in the active site(s) at the times when the splicing reactions occur. Genes encoding Prp8p have been isolated from yeast and other eukaryotes and are extraordinarily well conserved, consistent with the multiple roles of this factor.

This thesis focuses on the only region of yeast Prp8p not known to be common to other eukaryotes: a repetitive acidic and proline-rich domain at the N-terminus. Removal of part or all of this domain inhibits function, but cells lacking this domain are viable if the truncated protein is overproduced. A reconstruction approach suggests that proline is the most important feature of this domain, and thus function may be analogous to proline-rich regions of other proteins which in general promote complex assembly. The phenotype of truncation mutants is consistent with this. Spliceosome components are present in the yeast nucleus at lower concentration than in other eukaryotes, suggesting a reason as to why this otherwise strongly conserved protein possesses the extra domain.

This thesis also describes the analysis of several mutants of yeast *PRP8* which have the highly unexpected phenotype of a block to cell cycle progression. The data indicate that these mutants also affect splicing. The growth defect of one of them (*dbf3-1*) is suppressed by a cDNA copy of the *TUB1* gene. *TUB1* contains an intron and encodes a microtubule subunit (α-tubulin) functional in M-phase. Suppression separates the cell cycle and splicing defects, as the splicing defect is unaffected by suppression. These data strongly support the hypothesis that the cell cycle defect is a secondary consequence of a splicing defect, the link being a gene (*TUB1*) which functions in the cell cycle and which contains an intron. The cell cycle block is

observed because the splicing defect is mild: except for the *TUB1* intron, splicing in a *dbf3-1* cell is sufficient to maintain growth.

Abbreviations

Amp	Ampicillin
APS	Ammonium persulphate
ARS	Autonomous Replication Sequence
ATP	Adenosine 5'-triphosphate
A_{260}	Absorbance at 260nm
BCIP	5-Bromo-4-chloro-3-indolyl Phosphate
bisacrylamide	*N,N'*-methylene-bisacrylamide
bp	Base pair
BPS	Branch Point Sequence
BSA	Bovine Serum Albumin
°C	Degrees celsius
CEN	Centromere
cDNA	Complementary DNA
Ci	Curie ($2.2x10^{12}$ disintegrations/min)
CTP	Cytidine 5'-triphosphate
dATP	Deoxyadenosine 5'-triphosphate
dCTP	Deoxycytidine 5'-triphosphate
DEPC	Diethylpyrocarbonate
dH_2O	Distilled water
DMSO	Dimethyl Sulphoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
ds	Double-stranded
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
g	Gram or Grammes
Galactose	D-galactose
Glucose	D-glucose
hr	Hour(s)
HEPES	*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulphonic acid
IPTG	Isopropyl-β-D-thiogalactopyranoside
k	Kilo- (10^3)
kb	Kilobase
kDa	KiloDalton
l	Litre(s)
LB	Luria-Bertani broth (medium)
M	Molar
m	Milli- (10^{-3})
mg	Milligram(mes)
m_3G	2,2,7-trimethyl guanosine
mM	Millimolar
min	Minute(s)
ml	Millilitre(s)

MOPS	3-(*N*-morpholino)propanesulphonic acid
mRNA	Messenger RNA
N	Purine or pyrimidine
NBT	Nitroblue tetrazolium
NP-40	Nonidet P-40 detergent
NTP	Nucleoside 5'-triphosphate
OD_{600}	Optical Density at 600nm
OLB	Oligo Labelling Buffer
ORF	Open Reading Frame
SDS-PAGE	SDS-Polyacrylamide Gel Electrophoresis
PAS	Protein A-Sepharose
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
Pol	Polymerase
R	Purine nucleotide
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	Ribonuclease inhibitor
rpm	Revolutions per minute
rRNA	Ribosomal RNA
SDS	Sodium dodecyl sulphate
sec	Second(s)
snRNA	Small Nuclear RNA
snRNP	Small Nuclear Ribonucleoprotein Particle
SSC	Saline sodium citrate buffer
TAE	Tris-acetate buffer
TBE	Tris-borate buffer
TBS	Tris-buffered saline
TE	Tris/EDTA
TEMED	N,N,N',N'-tetramethyl ethylenediamine
tRNA	Transfer RNA
Tris	Tris(hydroxymethyl)aminomethane
µ	Micro (10^{-6})
µg	Microgram(mes)
µl	Microlitre(s)
µM	Micromolar
UV	Ultraviolet light
v/v	Volume per unit volume
w/v	Weight per unit volume
X-Gal	5-Bromo-4-chloro-3-indolyl-β-D-galactoside
Y	Pyrimidine nucleotide
YCp	Yeast Centromeric Plasmid
YEp	Yeast Episomal Plasmid

Amino Acids (single letter code)

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

A Note on Nomenclature

Almost all yeast genes are designated with three letters and a number. In this thesis yeast genes are described in italics according to convention, with upper-case for dominant alleles (mostly the wild-type) and lower-case for recessive alleles. Alleles are numbered (*prp8-1*, *prp8-2*, etc). Names of deletion mutations incorporate the character Δ. Names for gene products are not italicised, and where there is no common name (eg tubulin, U5 snRNA) the gene product is named after the gene, thus *PRP8* encodes the Prp8p (p for protein). The adopted system of numbering within the *PRP8* gene assigns A of the ATG start codon as position 1.

CONTENTS

CHAPTER 8: ISOLATION OF SUPPRESSOR OF *dbf3-1* FROM A cDNA LIBRARY

Chapter 1.

Introduction.

1.1: The Discovery of Introns.

Between 1974 and 1977, a series of observations was made which could not be accommodated under the then-current paradigm, that in eukaryotes mRNA was copied directly from a gene, as it is in prokaryotes (reviewed in Witowski, 1988). One such observation was the direct visualisation by electron microscopy of hybrids between viral early transcripts and cDNA, which showed long RNA displacement loops. To account for this, Berget *et al.* (1977) proposed a mechanism of RNA splicing by which non-coding regions interrupting eukaryotic genes (introns) are excised from precursor transcripts by precise cleavage/ligation events which covalently join adjacent coding regions (exons) to form mature mRNA. Introns are now known to occur in the vast majority of eukaryotic genes, and their removal is an essential step in gene expression. The realisation of the existence of so-called split genes and splicing was revolutionary, and initiated massive effort to elucidate the mechanisms by which splicing reactions occur. For their pioneering work in this field, Richard Roberts and Philip Sharp were awarded the 1992 Nobel prize.

1.2: The Mechanism of Splicing.

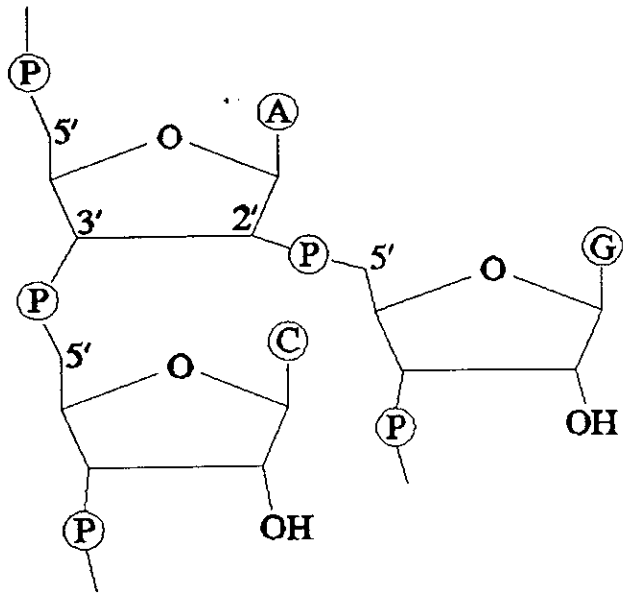
Description of the splicing mechanism followed the development of the use of extracts from HeLa cells to excise an intron *in vitro* from a synthetic RNA substrate (Hernandez and Keller, 1983; Krainer *et al.*, 1984). Branched RNA structures had been detected in HeLa nuclear RNA (Wallace and Edmonds, 1983), and these were shown to be intermediates and products of splicing, both *in vivo* (Domday *et al.*, 1984; Zeitlin and Efstradiatis, 1984; Rodriguez *et al.*, 1984) and in *in vitro* reactions (Padgett *et al.*, 1984; Ruskin *et al.*, 1984; Lin *et al.*, 1985). The discovery that a synthetic substrate was accurately cleaved at the boundary between upstream exon and intron (the 5' splice site)

Figure 1.1

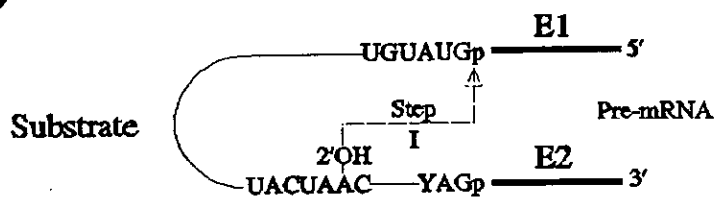
(A): Structure of the Branchpoint. After step I, three phosphodiester bonds impinge on the 2', 3', and 5' positions of the ribose at the branchpoint. P represents phosphate.

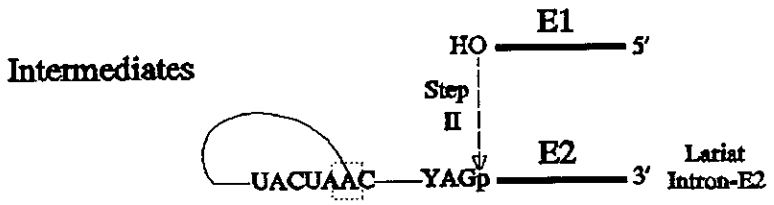
(B): The Mechanism of Splicing. The two transesterification reactions, step I and step II are shown taking place on a substrate with 5' and 3' splice site and branchpoint sequences corresponding to the consensus for yeast. The nucleophiles (hydroxyl groups) and the phosphates they attack (p) are indicated. The position of the branch, detailed in A above, is outlined (dashed box). The reaction joins the upstream and downstream exons (E1 and E2 respectively, thick lines), displacing the intron (thin line).

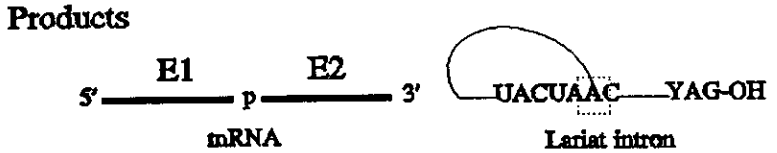
A


P
5′
O
A
3′
2′
P
5′
O
G
P
5′
O
C
P
OH
P
OH

B


Substrate
E1
UGUAUGp
5′
Step
I
Pre-mRNA
2′OH
E2
UACUAAC
YAGp
3′


Intermediates
E1
HO
5′
Step
II
E2
UACUAAC
YAGp
3′
Lariat
Intron-E2


Products
E1
E2
5′
p
3′
mRNA
UACUAAC
YAG-OH
Lariat intron

when incubated with a cell extract established that splicing occurred in two steps (Krainer *et al.*, 1984). Both of these steps involve transesterification reactions: nucleophilic attack of a ribose hydroxyl group on the phosphate at an exon/intron boundary (also known as phosphoryl transfer). In the first transesterification step, the 2' hydroxyl of an adenosine within the intron (the branchpoint) attacks the 5' splice site, displacing the upstream exon with a free 3' hydroxyl, and creating a branched intermediate with a lariat structure. The ribose at the lariat junction is connected via three phosphodiester bonds, impinging upon the 2', 3' and 5' positions. In the second transesterification step, the 3' hydroxyl of the upstream exon attacks the phosphate between the intron and the downstream exon (the 3' spice site) displacing the intron (as a lariat) and joining the two exons (see Figure 1.1).

Techniques pioneered by Moore and Sharp (1992) to introduce modified nucleotides at specific positions in synthetic precursors have confirmed this reaction pathway. By incorporating chiral phosphate analogues into the splice sites, the transesterification reactions were confirmed to proceed via single in-line SN2 nucleophilic displacement (Maschhoff and Padgett, 1993; Moore and Sharp, 1993). All nuclear introns are spliced by this common reaction mechanism, with the exception of introns in tRNA, which are spliced by endonuclease and ligase protein enzymes, in an unrelated reaction (for reviews of tRNA splicing, see Culbertson and Winey, 1989; Zillman *et al.*, 1991, Hopper and Martin, 1992).

1.3: The Origin of Introns.

Group II introns are found in precursor transcripts in bacteria and organelles of fungi, lower eukaryotes and plants and differ from nuclear pre-mRNA introns in that they are extensively conserved. For a subset of Group II introns, splicing can be catalysed *in vitro* by the intron itself, in a wholly RNA-catalysed reaction which does not require protein enzymes or energy input. Despite these differences, group II splicing proceeds by way of the same intermediates and products as nuclear pre-mRNA splicing suggesting a common mechanism (Cech, 1993). Self-splicing introns of this type are proposed to be representative of archetypal introns in a compelling theory of intron origin (Cavalier-Smith, 1985). According to this model, the first introns invaded the

genome of a progenitor eukaryote. Two observations of self-splicing introns suggest a means by which the archetypal introns could have spread through and between genomes. Firstly, group II introns are able to insert into an exon-exon boundary in mature mRNA, or into foreign RNA containing the same, short target sequence, in the reverse of the splicing reaction (Morl and Schmelzer, 1990). Secondly, certain group II introns contain open reading frames which could encode reverse transcriptase activity, necessary to introduce intron-containing RNAs into chromosomal DNA by recombination (Augustin *et al.*, 1990). Transposition of group II introns to ectopic sites *in vivo* has been recently demonstrated (Mueller *et al.*, 1993; Sellem *et al.*, 1993). The theory envisages that self-transposing introns gradually through evolution lost catalytic activity, which transferred to being *trans*-acting and encoded outside the intron. Plausible evolutionary intermediates in this process can be seen in modern-day group II trans-splicing systems (Sharp, 1991). The *psaA* gene, for example, is encoded by three trans-spliced exons scattered on both strands of the chloroplast genome. Splicing of *psaA* exons 1 and 2 requires assembly with a third trans-acting RNA, encoded by the chloroplast gene *tscA*, into a catalytic group II structure (Goldschmidt-Clermont *et al.*, 1991). A prediction of this model of intron origin is that all pre-mRNA intron splicing is RNA-catalysed. This notion is not yet supported by firm evidence, but has gained much credence in recent years.

An alternative view is that introns are far more ancient, having been present in a progenitor organism in which genes evolved as discontinuous structures. The primordial exons encoded stable protein-folding domains, and RNA splicing allowed the production of multi-domain proteins from exon modules (Darnell and Doolittle, 1986; Gilbert *et al.*, 1986; Sharp, 1995). Evidence in support of this model has come from an examination of intron position in evolutionarily ancient genes such as that encoding pyruvate kinase, an enzyme critical for energy metabolism. In chicken the coding sequence is interrupted by nine introns, several of which lie at boundaries between structural domains, as predicted by this model. Also, an intron in the pyruvate kinase nucleotide-binding fold occurs at the same position as an intron in the nucleotide-binding fold of alcohol dehydrogenase, supporting the model that the common progenitor from which these domains evolved also contained an intron (Lonberg and Gilbert, 1985). The proposition that the progenitor organism had a split

gene structure comparable to that in eukaryotes implies that complete intron loss has occurred in the recent evolution of modern day prokaryotes. The two models of intron origin are not totally mutually exclusive, and elements of both may be true.

1.4 The Study of Splicing.

A significant part of our understanding of pre-mRNA splicing has come from studies in the budding yeast *Saccharomyces cerevisiae*. This micro-organism (known henceforth in this thesis as yeast) is genetically tractable and has a simple genome, allowing factors involved in splicing to be identified and characterised genetically. Furthermore, as with HeLa cells, extracts from yeast accurately splice synthetic precursors *in vitro* (Lin *et al.*, 1985). More than 30 yeast genes encoding splicing factors have been isolated, either by screening randomly-generated collections of mutants for those affecting splicing, or by directed genetic approaches (reviewed in Beggs, 1995). Most of these genes are known by the three-letter code *PRP*, for **p**re-m**R**NA **p**rocessing. Splicing has also been studied in *Xenopus*, *Drosophila*, *Schizosaccharomyces pombe* and other fungi, monocotyledonous and dicotyledonous plants, and in human cell lines. Cultured mammalian cells have the advantage over yeast that splicing factors are far more abundant, facilitating their purification. Consensus in the nomenclature of mammalian splicing factors is lacking.

It has become clear that the splicing machinery in eukaryotes is strikingly well conserved, and that the process of splicing does not differ fundamentally between yeast and mammals. The subsequent chapters in this thesis are primarily concerned with yeast, but this review will deal with the yeast and mammalian systems together as in most cases their studies have been complementary.

1.5 The Distribution of Introns.

The number of introns in a pre-mRNA can vary from none to more than 50, and intron size varies from less than 10^2 nucleotides to more than 10^5. Yeast introns are somewhat unusual: they are small, generally less than 400 nucleotides, and none approaches the gigantic size of some introns of metazoa. Mammalian genes often

contain numerous introns, whereas yeast genes contain either one or none (with two exceptions, *MATα1* (Miller, 1984) and *RPL8A*, which each contain two introns). Yeast has a more stringent requirement for conserved *cis*-acting sequences adjacent to the splice sites and the branchpoint. As a consequence, mammalian *in vitro* splicing systems will accurately splice yeast or plant introns (Ruskin *et al.*, 1986; Brown *et al.*, 1986), but mammalian introns are in general not spliced by yeast systems (Beggs *et al.*, 1980; Langford *et al.*, 1983). Lastly, the frequency of introns in the yeast genome (estimated to be in 2.5% of non-ribosomal protein genes; Rodriguez-Medina and Rymond, 1994) is much lower than in other eukaryotes. Even in the fission yeast *Schizosaccharomyces pombe* about half of all genes contain introns (Potashkin *et al.*, 1989).

These differences are thought to be due to the potent homologous recombination system of yeast (Fink, 1987). Combined with reverse transcriptase activity, recombination in yeast would convert an intron-containing gene into a cDNA copy of a spliced mRNA derived from it, effectively deleting the intron. In other organisms with less potent systems of homologous recombination, cDNA copies of spliced genes would be more likely to be dispersed in the genome. Consistent with this, yeast have very few pseudogenes, whereas mammalian genomes have many. Thus the model proposes that a progenitor of yeast contained a similar distribution of introns to other eukaryotes, but that most of these have been deleted by gene conversion, through evolution (Fink, 1987). Yeast introns are generally located at the 5' ends of open reading frames, or in 5' untranslated regions, suggesting that these remaining introns are refractive to deletion by recombination.

A corollary of these differences is that in yeast, splicing as a regulatory step in gene expression is far less widespread than in metazoa, where programs of tissue- or sex-specific differentiation and development are often controlled by splicing events. Of the few examples of regulated splicing which are known in yeast, none involves the use of alternative splice sites.

1.6 SnRNPs and the Spliceosome: An Overview.

Unlike group II introns, the splicing of nuclear introns involves numerous *trans*-acting factors which assemble on an intron to form a complex called the spliceosome, in which the transesterification reactions take place. The primary components of the spliceosome are four ribonucleoprotein particles, called snRNPs (pronounced "snurps"). These contain small, metabolically stable RNA molecules (small nuclear RNAs, or snRNAs; Guthrie and Patterson, 1988; Lührmann *et al.*, 1990). The snRNPs are named according to the snRNA they contain, thus the U1, U2, U5 and U4/U6 particles contain respectively U1, U2, U5, and U4 plus U6 snRNA. In the U4/U6 snRNP, the U4 and U6 snRNAs are associated by extensive base-pairing. In each snRNP, the U1, U2, U4 and U5 snRNAs are complexed with a set of seven or eight common core proteins, and a variable number of snRNP-specific proteins. A host of other proteins (non-snRNP proteins) participate in spliceosome assembly and splicing, and may be stable spliceosomal components, or interact transiently with the spliceosome.

The assembly of the spliceosome is, then, the essence of nuclear pre-mRNA splicing: the snRNP particles and other proteins recognise sequence determinants within hnRNA transcribed by Pol II which specify an intron, a highly ordered pathway then proceeds to assemble a catalytic centre within a large complex, and transesterification takes place (for recent reviews of splicing, see Green, 1991; Moore *et al.*, 1993; Newman, 1994; Lamond, 1995). The snRNA components of snRNPs contribute, at least in part, to the structure of the spliceosome active site. ATP hydrolysis is required to drive spliceosome assembly, and hence nuclear pre-mRNA splicing is an energy requiring process, even though the total number of phosphodiester bonds is conserved. The process of spliceosome assembly is examined in detail below.

1.7 The *Cis*-acting Intron Determinants.

Nuclear pre-mRNA introns lack conserved sequences except for minimal elements at the splice sites and the branchpoint and between the branchpoint and the 3' splice site. This lack of conservation is remarkable in view of the huge size of some

metazoan introns and presents these organisms with the problem of selection to ensure splicing fidelity. Intron sequence determinants have been identified through sequence comparisons of both yeast and mammalian introns (Mount, 1982; Guthrie, 1986), and their contribution to function determined by mutation.

1.7.1 The 5' Splice Site.

The general consensus of the 5' splice site is AG/GURAGU (where / denotes the cleavage site; Mount, 1982). In yeast the 5' splice site is more highly conserved (consensus R/GUAUGU, Rymond and Rosbash, 1992). This sequence is required for splicing (Green, 1986). Mutations within this sequence generally abolish splicing in yeast (Vijayraghavan *et al.*, 1986; Fouser and Friesen, 1986), or lead to the use of local, alternative (cryptic) sites in mammalian systems (Aebi, 1986 and 1987). Mutation of the invariant GU nucleotides often does not prevent lariat formation but prevents the later steps, causing the accumulation of "dead-end" lariat intermediates (Newman *et al.*, 1985; Fouser and Friesen, 1986). Clearly, step I and step II have different sequence requirements. Instances are known of point mutations in 5' splice sites (or 3' splice sites) in introns of genes such as β-globin, RB (retinoblastoma) or CFTR (cystic fibrosis transmembrane conductance regulator) causing human genetic disease due to exon-skipping or cryptic splice site activation (Antoniou, 1995).

1.7.2 The Branchpoint.

Yeast introns contain the sequence UACUAAC (the branchpoint sequence, BPS) which is usually located 20-60 nucleotides upstream of the 3' splice site, and which is required for splicing (Pikielny *et al.*, 1983). The 3'-most adenosine in this sequence (shown in bold) provides the nucleophile in step I and is the site of the lariat branch. The sequence is almost completely conserved (Rymond and Rosbash, 1992). Mutations within this sequence generally abolish splicing, although as with the 5' splice site, some mutations allow step I but not step II, resulting in the accumulation of lariat intermediates (Fouser and Friesen, 1986; Jacquier and Rosbach, 1986; Vijayraghavan *et al.*, 1986). That mutations in cis-acting sequence have this phenotype is suggestive of a proofreading mechanism, operating to ensure high-fidelity splicing.

Mammalian introns contain a similar, although less strictly conserved BPS, of consensus UNCURAC (Rymond and Rosbach, 1992). Mutational analysis however suggests a far less stringent requirement for a sequence in mammalian introns conforming to this consensus than in yeast: when the β-globin BPS was deleted, multiple cryptic sites were activated, although lariat formation always occurred at an adenosine, located 22-37 nucleotides upstream of the 3' splice site (Ruskin *et al.*, 1985). Competition between duplicated branchpoint sequences in introns which otherwise lack cryptic sites shows that base substitutions or deletions within this sequence abolish splicing *in vitro*, and promote use of competing sites *in vivo*, therefore the mammalian BPS makes a significant contribution to splicing efficiency (Reed and Maniatis, 1988).

1.7.3: The 3' Splice Site and Polypyrimidine Tract.

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

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

In yeast introns, the branch site is specified primarily by the UACUAAC sequence independent of pyrimidine-richness in the 3' region. A model substrate truncated immediately after the BPS underwent step I when ligated via its 3' end to a homopolymeric oligoribonucleotide of any sequence. The only requirement for lariat formation was intron length (at least 29 nucleotides following the BPS; Rymond *et al.*, 1987). The yeast polypyrimidine tract, though not strictly requisite, greatly enhances the efficient usage of an adjacent AG dinucleotide in alternative 3' splice site competition assays (Patterson and Guthrie, 1991). Moreover, distal pyrimidine-rich sites can out-compete proximal purine-rich sites (Patterson and Guthrie, 1991), so scanning cannot be the sole mechanism for 3' splice site selection in this system.

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

1.8 SnRNAs and SnRNPs.

SnRNPs are highly abundant in mammalian cells (10^5 to 10^6 copies per nucleus; Lührmann *et al.*, 1990), but far less so in yeast (approximately 10^2 copies per nucleus; Wise *et al.*, 1983; Reidel *et al.*, 1986) reflecting the relative rarity of yeast introns. The core proteins of mammalian snRNPs are designated B, B', D_1, D_2, D_3, E, F, G, in order of decreasing apparent molecular weight (B and B' are apparently splice variants encoded by the same gene). They associate with a single-stranded motif present in U1, U2, U4 and U5 snRNAs (Mattaj, 1986). This sequence (the Sm-site) is conserved in both yeast and mammalian snRNAs (Riedel *et al.*, 1987), and has the consensus $RAU_{4-6}GR$ flanked by a double-stranded stem. The core proteins are recognised by sera (so-called anti-Sm sera after the patient Smith) from patients suffering from the autoimmune disease systemic lupus erythematosus. Anti-Sm sera also immunoprecipitate yeast snRNAs (Silicano *et al.*, 1987; Tollervey and Mattaj, 1987), showing that immunological determinants of the core proteins are conserved.

Yeast genes encoding D1 and D3 core proteins have sequence similarity to their higher eukaryote homologues, and have been identified fortuitously from their locality to other genes (Rymond, 1993; Lehmeier *et al.*, 1994; Roy *et al.*, 1995). A burgeoning number of genes are now known with homology to the core proteins, and sequence comparisons amongst these have enabled the identification of a shared motif (the Sm-motif; Hermann *et al.*, 1995; Cooper *et al.*, 1995; Séraphin, 1995)

SnRNAs vary with respect to size and sequence, but their secondary structures are highly conserved (Guthrie and Patterson, 1988). The U1, U2, U4 and U5 snRNAs are transcribed by RNA polymerase II, and have the cap structure 2,2,7-trimethylguanosine (m_3G). In contrast, U6 snRNA is transcribed by RNA polymerase III, and has a γ-monomethyl phosphate cap (Parry *et al.*, 1989; Singh and Reddy, 1989). A remarkable discovery of recent years has been the finding that U6 snRNA, which lacks the Sm-site, is specifically associated with proteins homologous to the common core proteins suggesting the existence of a U6-specific core complex. These proteins have been named Uss, for U-six-snRNP (Uss1p, Cooper *et al.*, 1995; Uss2p, Séraphin, 1995)

The biogenesis of snRNPs U1 to U5 requires transport in both directions across the nuclear envelope. The snRNAs are exported posttransciptionally to the cytoplasm, where the snRNP particle is formed (Mattaj, 1988). In the cytoplasm, the core proteins assemble onto the Sm-site to form the core complex. Thereafter, and requisite upon prior core complex formation, the m^7G cap is hypermethylated to m_3G by a cytoplasmic methyltransferase activity (Plessel, 1994). SnRNP-specific (non-core) protein assembly, and import of the snRNP complex into the nucleus via nuclear pores, then proceed (Nelissen *et al.*, 1994). Nuclear import of snRNPs is directed by a bipartite nuclear localization signal (NLS), consisting of the m_3G cap, and protein motifs presented by the core complex (Hamm *et al.*, 1990; Fischer *et al.*, 1993).

1.9 Spliceosome Assembly.

In the nucleus, nascent RNA polymerase II transcripts associate with more than 20 different proteins to form heterogeneous nuclear ribonucleoprotein (hnRNP) particles (Dreyfuss, 1993). RNA added to splicing-competent cell extracts is assembled

into equivalent complexes (Konarska and Sharp, 1986). That this complex is not specific to the splicing reaction is apparent because its assembly occurs independently of splice site sequences within the RNA. Also, competent splicing substrates assembled into hnRNP cannot be chased into spliced products in the presence of competitor pre-mRNA (Michaud and Reed, 1991). However, although the hnRNP proteins do not commit a bound RNA to the splicing pathway, they do exert considerable influence on subsequent spliceosome assembly (Swanson, 1995). Different splicing substrates bind different subsets of hnRNP proteins in vitro (Bennet *et al.*, 1992a). The SF5 factor, which was purified on the basis of its activity required for alternative splicing of a model substrate, turned out to be the hnRNP A1 protein (Mayeda and Kainer, 1992), implicating hnRNP proteins in alternative splicing. It is an intriguing possibility that genes implicated in human disease, and encoding homologues of hnRNP proteins (see for example Ma *et al.*, 1993) may be alternative splicing factors.

1.9.1: Association of the U1 SnRNP.

The first step in the formation of the spliceosome is the binding of the U1 snRNP to the 5' splice site (Mount *et al.*, 1983). U1 snRNA has long been known to include at its 5' end the complement of the 5' splice site consensus (Lerner *et al.*, 1980; Rogers and Wall, 1980); this was the first clue that snRNPs might function in splicing. This region of U1 snRNA is essential for splicing *in vitro* (Krämer *et al.*, 1984) implying that Watson-Crick base-pairing interactions are involved in U1 snRNP binding. This was confirmed by the demonstration that mutations within these regions could be rescued by compensatory mutations restoring complementarity (Zhuang and Weiner, 1986; Séraphin *et al.*, 1988; Silicano and Guthrie, 1988; Nelson and Green, 1990). Also, in competition assays for use of alternative 5' splice sites, the site used is under certain conditions the one most complementary to U1 snRNA (Nelson and Green, 1990).

U1 snRNP binding to pre-mRNA occurs in the absence of ATP and U2 snRNP (Legrain *et al.*, 1988; Ruby and Abelson, 1988; Séraphin and Rosbach, 1989), whereas the interaction of U2 snRNP with pre-mRNA requires ATP and prior interaction with U1 snRNP (Seraphin *et al.*, 1988, Ruby and Abelson, 1988; Barabino *et al.*, 1990). There is thus a requisite order of snRNP addition in spliceosome formation, starting

with U1. In the yeast system, two U1 snRNP/pre-mRNA complexes can be resolved by native gel electrophoresis. In the first of these, called CC1 (commitment complex 1), the U1 snRNP interacts with pre-mRNA via the 5' splice site; this is the only pre-mRNA sequence element required for CC1 formation. The second complex, CC2, represents an additional interaction, in which the 3' end of the intron is recognised by protein-RNA interactions. From the order of assembly of these complexes, and the kinetics of formation of a similar commitment complex in the mammalian system (called E complex; Michaud and Reed, 1993), an interaction between the components recognising the 3' end of the intron and the U1 snRNP has been proposed.

One of the factors involved in recognising the 3' end of the intron is U2 snRNP Auxiliary Factor (U2AF). This factor was purified from HeLa cells as an activity required for the formation of a U2 snRNP/pre-mRNA complex (Zamore and Green, 1989). U2AF is a heterodimer of 35 kD and 65 kD components ($U2AF^{65}$ and $U2AF^{35}$ respectively). $U2AF^{65}$ is essential for splicing, and can be cross-linked to the polypyrimidine tract by UV light, suggesting a direct physical contact (Zamore and Green, 1989). $U2AF^{65}$ binds RNA via three RNP domains, a common amino acid sequence motif amongst RNA binding proteins (Zamore *et al.*, 1992). The preferred RNA sequence bound by $U2AF^{65}$ has been determined by iterative *in vitro* genetic selection from a randomly generated pool. In this method, also known as the Selex procedure, oligonucleotides degenerate for all four nucleotides at multiple positions are transcribed *in vitro* to generate a pool of RNA. In the selection process, the RNA pool is combined with immobilised, purified protein (in this case $U2AF^{65}$). The subset of the pool which binds to the protein is then reverse transcribed and amplified by the polymerase chain reaction (RT-PCR), and the process repeated. Those sequences remaining after multiple rounds of binding and amplification are bound with an affinity several orders of magnitude greater than the original pool, and represent the optimal ligand sequence. With $U2AF^{65}$ this procedure selects degenerate, uridine-rich sequences, supporting the view that it is a general splicing factor, able to bind to the polypyrimidine tracts of the majority of mammalian introns (Singh *et al.*, 1995).

The events which associate the U1 snRNP bound at the 5' splice site, and proteins such as U2AF bound at the polypyrimidine tract are of crucial importance because they are the basis of exon pairing, the process which in higher eukaryotes

determines which 5' spice site is spliced to which 3' splice site. No direct physical interactions between $U2AF^{65}$ and the U1 snRNP are known, rather the two seem to be linked by members of the SR protein family, which act as bridging factors. Two SR proteins in particular, SC35 and SF2/ASF, interact physically with an integral protein component of the U1 snRNP, U1-70K, and also with the $U2AF^{35}$ subunit of U2AF (Wu and Maniatis, 1993; Kohtz *et al.*, 1994).

The yeast homologue of $U2AF^{65}$ is Mud2p. The *mud* (*mutant-u-die*) mutants were identified in a screen for factors interacting with U1 snRNA. The *mud* mutations alone have no detectable phenotype, but in combination with mutations in U1 snRNA are synthetic lethal (Liao *et al.*, 1993). Mud1p is the yeast homologue of U1-A, a U1 snRNP integral protein. Mud2p exhibits 49% similarity and 31% identity to $U2AF^{65}$, and is present in commitment complexes in contact with the pre-mRNA (Abovich *et al.*, 1994). The binding of Mud2p to pre-mRNA requires the BPS (UACUAAC) whereas $U2AF^{65}$ requires only pyrimidine-rich sequence (Singh *et al.*, 1995). This is likely to be the basis of the differing requirements of the yeast and mammalian systems for intron 3' sequence elements, at least in part. One surprising finding is that the *MUD2* gene is dispensable in yeast. *Drosophila* and *S. pombe* U2AF homologues are essential (Abovich *et al.*, 1994). Perhaps in yeast other compensating functions identify the BPS. It has been suggested that the greater sequence complementarity between the U2 snRNA and the BPS (which basepair during spliceosome assembly) in yeast reduces the requirement for a U2AF targeting function (Hodges and Beggs, 1994).

1.9.2 Commitment and Alternative Splicing.

The U1 snRNP/pre-mRNA complexes described above form in the absence of ATP, nevertheless they are not exchangeable - incubation of these complexes with excess competitor pre-mRNA does not prevent chase into spliceosomes and spliced RNA products. This is the functional definition of commitment. It is thus apparent that commitment complex formation and the exon pairing events must be primary points of modulation in the regulation of alternative splicing.

The enormous variety of alternative splicing patterns which have been described seem to be based on just four modes of regulation: the intrinsic strength of the splice site sequence, RNA secondary structure, the concentrations of general

splicing factors, and *trans*-acting regulators of splicing (Valcárcel *et al.*, 1995). The yeast intron-containing gene *MER2* encodes a protein essential for meiotic recombination and which is only expressed during meiosis. The 5' splice site of *MER2* differs from the consensus and is poorly recognised by the U1 snRNP and not normally spliced. During meiosis the *trans*-acting factor Mer1p augments the 5' splice site interaction with U1, allowing *MER2* splicing and expression (Engebrecht *et al.*, 1991; Nandabalan *et al.*, 1993). Mer1p bears homology to hnRNP K (the KH motif, a putative RNA recognition motif; Siomi *et al.*, 1993) and may contact RNA directly.

An example again from yeast of an RNA secondary structure affecting splicing is the negative feedback which regulates expression of ribosomal protein (rp) gene *L32*. Pre-mRNA from this gene folds into a secondary structure which is bound and presumably stabilised by rpL32 protein. This structure obscures part of the 5' splice site. The U1 snRNP still binds the rpL32/pre-mRNA complex, but splicing is inhibited (Daveba *et al.*, 1986; Eng and Warner, 1991; Vilardell and Warner, 1994).

The SR family of proteins are essential for splicing (see, for example Fu and Maniatis, 1990) and are implicated as factors whose concentration influences the use of alternative splice sites. The six known members of this family share one or two RNP RNA recognition motifs in the N-terminus and a C-terminal region rich in serine-arginine dipeptides (the SR domain; Zahler *et al.*, 1992). They are conserved between mammals and *C.elegans*, though none are known in yeast. Each SR protein is able to restore activity to a splicing deficient S100 extract from mammalian cells, suggesting they have related functions. The first isolated SR protein was ASF, purified as an activity from adenovirus-infected 293 cells which promoted a non-default splicing pathway of the alternatively spliced SV40 early transcript (Ge and Manley, 1990). ASF was also purified as an essential splicing factor (SF2). *In vitro*, higher concentrations of this factor promote use of proximal sites in competitions of both 3' and 5' splice sites. This activity is counterbalanced by hnRNP A1 which promotes distal site use; the ratio of hnRNP A1 to SF2/ASF determines splice site choice (Mayeda and Krainer, 1992). ASF/SF2 is able to interact with factors bound at both the 5' and 3' regions of the intron as mentioned above, and presumably the role of SR proteins in alternative splicing is related to this.

Other features of the SR protein family are consistent with function as regulators of alternative splicing. They exhibit tissue-specific variation in their relative levels (Zahler *et al.*, 1993). Also, the SR domain is phosphorylated by a snRNP-associated kinase which also phosphorylates U1-70K (Woppmann *et al.*, 1993). It is possible that these phosphorylation events are the same as those observed by Mermoud *et al.*, who identified rounds of reversible phosphorylation occurring during the splicing process which may have a regulatory role (Mermoud *et al.*, 1992 and 1994). Different members of the SR family seem to have distinct roles in the regulation of alternative splicing *in vitro* (Zahler *et al.*, 1993). Fu (1993) showed that binding of the SR protein SC35 to a pre-mRNA was sufficient to commit it to splicing, and that different SR proteins committed different pre-mRNAs to the splicing pathway with differing potencies.

An example of a trans-acting specific regulator of splicing (as opposed to an essential splicing factor whose concentration influences splice site choice) is the *Drosophila* protein Sex-lethal (Sxl) which triggers female somatic sexual development (Cline, 1984). Sxl promotes splicing of an intron in the transformer (tra) gene to a distal 3' splice site. In the non-sex-specific program, tra is spliced to an alternative proximal site to producing a truncated protein. Sxl binds with high affinity to the polypyrimidine tract adjacent to the non-sex-specific 3' splice site and antagonises U2AF binding, and the distal site is utilised producing an mRNA encoding the full-length protein (Sosnowski *et al.*, 1989; Inoue *et al.*, 1990; Valcárcel *et al.*, 1993). Sxl lacks an SR domain, but if a chimera is created between Sxl and the SR domain of U2AF use of the non-sex-specific site is stimulated instead of repressed (Valcárcel *et al.*, 1993).

1.9.3 The Association of the U2 SnRNP.

The U1 snRNP/pre-mRNA complex directs U2 snRNP addition, the first ATP-dependent step of spliceosome assembly. The U1/U2/pre-mRNA complex is termed the prespliceosome. The association between the U2 snRNP and pre-mRNA involves the formation of another RNA-RNA duplex, between the BPS and an internal region of U2 snRNA (Parker *et al.*, 1987; Wu and Manley, 1989; Zhuang and Weiner, 1989). As predicted from the respective sequences, the branch site adenosine is bulged out of this duplex (Query *et al.*, 1994). The molecular architecture of the U2 snRNP has become

familiar from the extensive application of biochemical fractionation and yeast genetic techniques. At least eleven integral proteins are unique to the U2 snRNP (reviewed in Hodges and Beggs, 1994). Three of these copurify in the SF3a complex, and six in the SF3b complex (Behrens *et al.*, 1993; Brosi *et al.*, 1993a and b). On the basis of amino acid sequence similarity, three associated SF3a subunits, SAP61, SAP62 and SAP114 (nomenclature of Bennet *et al.*, 1992b) are the homologues of the yeast proteins Prp9, Prp11, and Prp21 (Bennet and Reed, 1993; Hodges and Beggs, 1994).

Mutations in the yeast genes encoding these three proteins and also Prp5p were originally identified in genetic screens (Hartwell, 1970; Vijayraghavan, 1989). When any pair of the conditional mutations *prp5-1*, *prp9-1*, *prp11-1* and *prp21-1* is combined in the same yeast strain, or when one of these alleles is combined with mutations in U2 snRNA, the double mutant is not viable, or more severely effected by temperature than either mutant alone (Ruby *et al.*, 1993). *spp91-1* is an allele of *PRP21* which was identified as a suppressor of *prp9-1* (Chapon and Legrain, 1992). Overexpression of wild-type *PRP9* or *PRP11* also has a suppressing effect on certain conditional alleles in this group (Ruby *et al.*, 1993). Thus both synthetic lethality and suppression suggest that members of this group are functionally related.

The *in vitro* splicing activity of cell extracts from *prp9-1* strains is heat-labile, and heat-inactivated extracts are blocked at the stage of U2 snRNP addition to the spliceosome (Abovich *et al.*, 1990). Similarly, inactivating other proteins in this group, either by heat treatment of a conditional mutant, or antibody binding to wild-type protein, inhibits the same step of the *in vitro* reaction (Ruby *et al.*, 1993; Legrain *et al.*, 1993; Arenas and Abelson, 1993). Direct physical interactions amongst this group have been examined by the two-hybrid assay (Chien *et al.*, 1991). In these experiments, peptides are fused to DNA-binding and activation domains of transcription factors, and expressed in pairwise combinations in yeast. When the peptides under test bind to each other with high affinity, the chimeras associate to form a hybrid transcription factor which activates expression of a reporter gene, or a selectable marker. This analysis showed that Prp9p forms a homodimer mediated by a central zinc-finger motif (CH1). Prp9p also binds Prp21p, and this interaction is inhibited by the conditional *prp9* mutations. Prp9p has an N-terminal leucine zipper-like motif, and a second, C-terminal zinc-finger (CH2) which is known to have an important function because

overproduction of Prp9p truncated of this domain is toxic in *prp9* yeast (Legrain *et al.*, 1993). The same dominant negative phenotype is generated when the CH domain of Prp11p is deleted (Legrain and Chapon, 1993). The Prp11p/Prp21p, and the Prp11p/Prp9p combinations in the two-hybrid assay also interact, but in the latter case only when Prp21p is also overexpressed, suggesting that Prp9p and Prp11p do not bind directly but require a molecular bridge, provided by Prp21p (Legrain and Chapon, 1993).

How does the U2 snRNP associate with the commitment complex? One possibility is that proteins of the SF3a complex associate with U2AF. An interaction has been demonstrated in the yeast between the $U2AF^{65}$ homologue Mud2p and Prp11p by the two-hybrid assay, and confirmed genetically by the observation that the *MUD2* gene is dispensable in wild-type but not in *prp11* mutant yeast (Abovich *et al.*, 1994). $U2AF^{65}$ is much less readily detected in spliceosomal complexes after addition of the U2 snRNP, suggesting that entry of U2 into the spliceosome displaces U2AF from the pre-mRNA (Bennet *et al.*, 1992a).

No mammalian homologue of Prp5p has yet been found. This protein bears motifs characteristic of the DEAD-box family of RNA-dependent ATPases. Several members of this family have demonstrated RNA helicase activities, and hence it is speculated that this protein catalyses the unwinding, or other rearrangement, of an RNA/RNA duplex, and that this rearrangement is required for U2 snRNP association (Hodges and Beggs, 1994). Precisely what is the substrate of this putative helicase is unknown, but U1 and U2 snRNAs are obvious candidates.

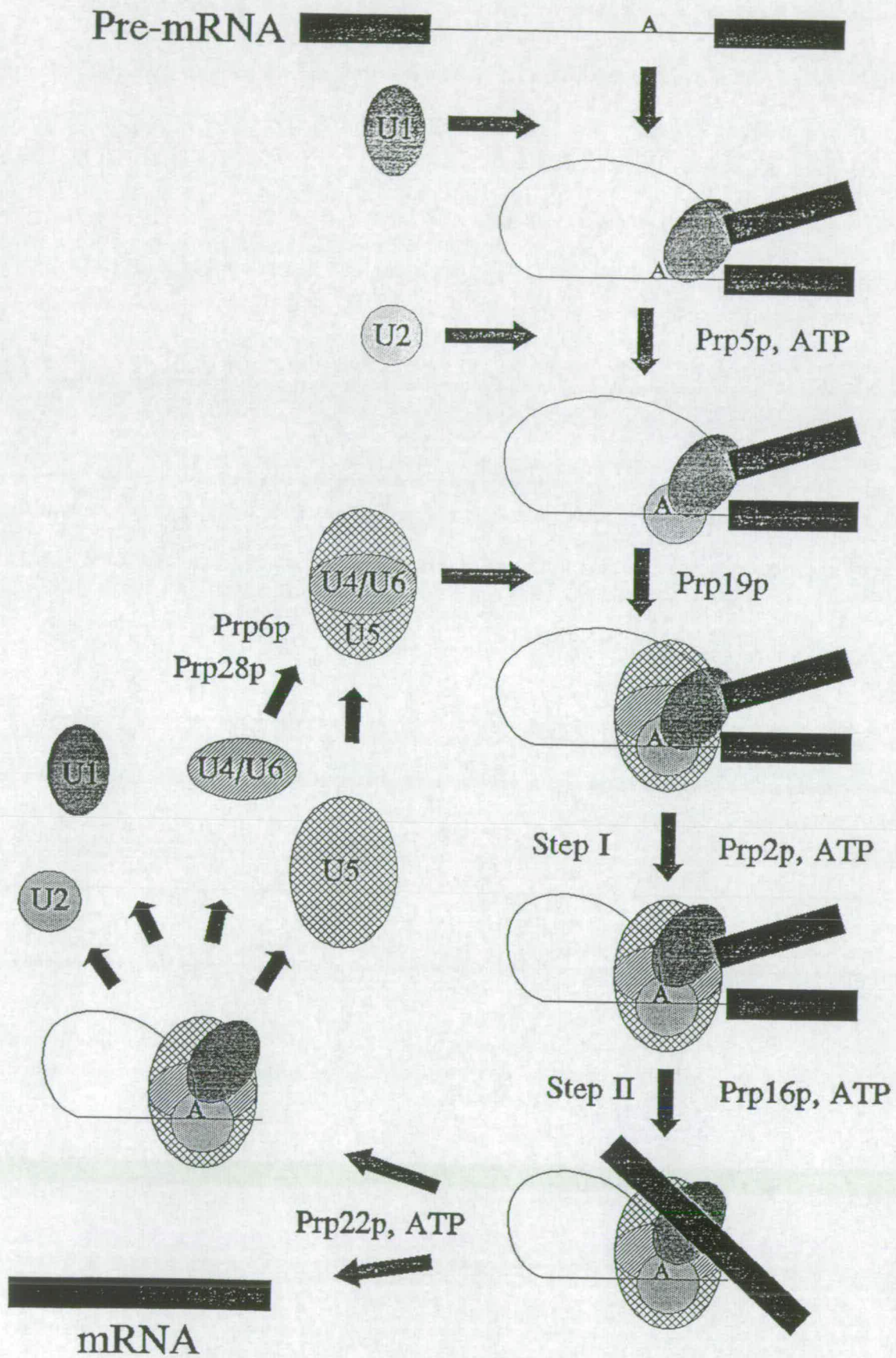
1.9.4 Association of the U4/U6.U5 Tri-snRNP Particle.

In the third and final step of snRNP addition to the spliceosome the U4/U6 and U5 snRNPs associate with the prespliceosome. It is likely that these snRNPs bind in the form of a pre-assembled unit termed the tri-snRNP (as it contains three snRNAs) which can be detected in cell extracts. The evidence to support this is that depletion of integral components of the U5 snRNP or the U4/U6 snRNP prevents either snRNP from entering the spliceosome. This has been observed with Prp8p (Brown and Beggs, 1992), U5 snRNA (Séraphin *et al.*, 1991), and Uss1p (Cooper *et al.*, 1995) in yeast, and in the mammalian system with U5 snRNA (Lamm *et al.*, 1991). Also, factors have been

identified which are present only in the tri-snRNP and later spliceosomal particles, and which are required for tri-snRNP assembly and for tri-snRNP association with the prespliceosome (Behrens and Lührmann, 1991; Utans *et al.*, 1992).

Figure 1.2: The Spliceosome Cycle.

SnRNP particles recognise intron determinants and assemble into a spliceosome in a defined, stepwise order. The steps in this pathway which are known to require ATP are indicated. Also shown are seven non-snRNP proteins which act at particular points. Five of these factors (Prp2p, Prp16p, Prp22, Prp5p, and Prp28p) are members of the DEAD/H-box family of putative RNA helicases. Adapted from Beggs (1995). Key: intron, thin line; exons, thick lines; branchpoint nucleotide, A.


Pre-mRNA
A
U1
A
U2
Prp5p, ATP
A
U4/U6
U5
Prp19p
Prp6p
Prp28p
A
U1
U4/U6
Step I
Prp2p, ATP
U5
U2
A
Step II
Prp16p, ATP
A
Prp22p, ATP
mRNA

1.10 Conformational Change in the Spliceosome.

After the four snRNPs have assembled on a pre-mRNA substrate, a series of conformational changes proceeds, culminating in the construction of an active site and the catalysis of transesterification. The underlying theme is of a dynamic network of RNA-RNA interactions, promoted by protein:protein and protein:RNA interactions. The temporal order of some of these conformational changes is still in question.

1.10.1 Disassociation of U4 and U6 SnRNAs.

The non-denaturing gel system of Pikielny *et al.* (1986) resolves three spliceosomal complexes. In their order of assembly, they are the prespliceosome (complex III), a slower-migrating complex (complex I) which corresponds to the prespliceosome plus the tri-snRNP particle, and a third complex (complex II), in which the intermediates and products of splicing can be detected. Complex II has an intermediate mobility between complex III and I, suggesting that the I to II transition is accompanied by a gross conformational change. U4 snRNA is readily detected in complexes I and III, but not II (Pikielny *et al.*, 1986; Cheng and Abelson, 1987; Parker *et al.*, 1987). Indeed, the U4 snRNP is dispensable for reactions in the spliceosome, and is not required beyond spliceosome assembly (Yean and Lin, 1991). However, an association between U4 snRNA and spliced products can be detected using biotinylated oligonucleotide probes (Blencowe *et al.*, 1989).

A mutant of yeast U4 snRNA (G13C) within the region which base-pairs to U6 and which is therefore expected to destabilise the U4/U6 interaction has a cold-sensitive phenotype, suggesting it causes hyperstabilization (Shannon and Guthrie, 1991). The implication is that the U4/U6 duplex exists in equilibrium with a hypothetical competing complex, and that G14C disturbs this equilibrium in favour of the second complex. According to this model, second-site mutations which destabilised the competing complex would suppress G14C. The compensatory change in U6 snRNA which restores complementarity to U4 was isolated as a dominant suppressor. Amongst recessive suppressors were U6 mutants outwith the region of U4/U6 duplex, and also mutations in the previously isolated *PRP24* gene. U6 snRNA is coprecipitated

by anti-Prp24p antibodies from wild-type extracts, and from extracts of the U4-G14C mutant, but not from extracts from cells bearing the *prp24* or U6 recessive suppressor alleles. This is consistent with the competing complex model, and suggests that after unwinding from U4, U6 binds to Prp24p. Three of the U6 mutations which suppress G14C are clustered at positions 38-42, and this is most likely to be the region of U6 bound by Prp24p. The sequence of Prp24p includes a RNP RNA recognition motif, and each of the *prp24* suppressor mutations changes a conserved amino acid in the RNP motif. *PRP24* also interacts genetically with *PRP28*: the combination of *prp24-1* and *prp28-1* conditional alleles is synthetic lethal. Prp28p is a member of the DEAD-box family of putative RNA helicases, and its function may be to unwind the U4/U6 duplex (Strauss and Guthrie, 1991). Taken together, these results suggest that the U4/U6 duplex as it exists in the U4/U6 snRNP is unwound at about the time the tri-snRNP associates with the prespliceosome. U6 snRNA remains tightly associated with the spliceosome; U4 snRNA has no further function and is only loosely associated with the spliceosome after the unwinding event.

1.10.2 Interactions between U5 snRNA and the Spliceosome.

Newman and Norman (1991) changed a G residue to an A at the invariant intron position 1 in an intron:β-galactosidase reporter construct. This defective splicing substrate does not express β-galactosidase because step II is blocked. A *trans*-acting suppressor was obtained which restored β-galactosidase expression by splicing not to the authentic (G to A) 5' splice site but to a cryptic site 12 nucleotides upstream. The suppressor mutation was in U5 snRNA gene, in a region encoding the loop of a phylogenetically conserved stem-loop structure (loop I). The base change was U to C at position 6 of this loop. These workers then subjected this region of U5 snRNA to directed mutagenesis, and obtained further base-changes which activate cryptic 5' splice sites. In each case loop positions 5 and 6 were complementary to positions -2 and -3, upstream of the step I cleavage site (Newman and Norman, 1992).

Mutation of the G residue of the 3' splice site AG also blocks step II and causes lariat intermediates to accumulate. These mutations are also suppressed by U5 snRNA loop I base changes, which in this case restore efficient splicing at the authentic 3' splice site. In this case suppression occurs when the first two positions of the

downstream exon are complementary to loop I positions 3 and 4. This suggests that, in the spliceosome, loop I of U5 snRNA interacts with regions of exon directly adjacent to both splice sites (Newman and Norman, 1992).

An analysis of these interactions as they occur in the mammalian spliceosome has been informative of their temporal order. Wyatt *et al.* (1992) incorporated a 4-thiouridine at the -2 position upstream of the 5' splice site in a synthetic substrate, by the ligation technique of Moore and Sharp (1992). When this substrate is assembled into spliceosomes and exposed to ultraviolet light, the 4-thio group is photoactivated and reacts to form a covalent crosslink with any molecule in the vicinity (approximately one bond-length away). The -2 position crosslinks to position 5 of U5 loop I (as expected from Newman's work) and also to loop I position 7. These interactions peak and start to diminish before the first tranesterification reaction occurs. As these contacts to the -2 position dissolve, contacts are formed between U5 and the adjacent nucleotide, position -1 upstream of the 5' splice site (Sontheimer and Steitz, 1992). These can be detected before step I, and persist until after step I. The interaction between U5 loop I (positions 3 and 4) and position 1 of the downstream exon is initiated after step I and persists until after step II (Sontheimer and Steitz, 1993). Contact has also been detected between U5 snRNA and the lariat-intron product in the region downstream of the 5' splice site (Wasserman and Steitz, 1992). Pre-mRNA positions with which an snRNA basepairs are expected to be conserved, but the exon positions with which U5 interacts are not conserved. Two explanations for this can be proposed: firstly, uridine has been envisaged as a relatively promiscuous partner, able to form base-pairing interactions with nucleotides other than A, its formal Watson-Crick partner. This may explain the predominance of uridines in U5 loop I. Secondly, it is possible that the interaction between (wild-type) U5 snRNA and exon positions is not strictly a base-pair at all.

These experiments detected two other notable interactions. The first was between position -2 upstream of the 5' splice site and protein p220, which is the mammalian homologue of yeast Prp8p (Wyatt *et al.*, 1992). The second was between intron position 2 and U6 snRNA; this interaction was initiated after step I and persists after step II (Sontheimer and Steitz, 1992).

Frank *et al.* (1992) screened for genetically interacting factors which were synthetic lethal with conditional mutations in the U5 snRNA loop I region. They

isolated conditional alleles of the previously identified *PRP17* gene, and the novel genes *SLU1*, *SLU2*, *SLU4* and *SLU7* (for *synthetic lethal with U snRNA*). The group of factors Slu4, Slu7, Prp16 and Prp18 (Vijayraghavan and Abelson, 1990) form a genetically interacting set required for step II. With the exception of Prp16p, these factors are candidates for integral components of the U5 snRNP. Slu7p includes a motif related to the zinc-finger (the zinc-knuckle) and implicated in RNA-binding. Its function appears to be to promote step II in introns with a long BPS to 3' splice site distance in a sequence non-specific fashion.

1.10.3 Interaction of U6 and U2 SnRNA.

Two regions of U6 become base-paired to U2 in the spliceosome. The 5' end of U2 base-pairs with a region at the 3' end of U6 to form helix II (Hausner *et al.*, 1990; Datta and Weiner, 1991; Wu and Manley, 1991). This region of U6 is single-stranded in the U4/U6 snRNP. The equivalent positions in yeast U2 and U6 can be altered without affecting splicing, so helix II may be of lesser importance in this system (Fabrizio *et al.*, 1989; Madhani *et al.*, 1990).

The second U2/U6 duplex (helix I) involves a region of U6 which is base-paired to U4 snRNA in the U4/U6 snRNP, thus helix I can only form after U4/U6 unwinding. This region of U6 had previously been assigned functions beyond simply base-pairing with U4, because mutations in this region of U6 cannot be suppressed by compensatory mutations in U4 which restore complementarity (Madhani *et al.*, 1990; Fabrizio and Abelson, 1990; Vankan *et al.*, 1990). The regions of U2 snRNA forming helices I and II are associated together in the U2 snRNP in an internal stem-loop structure, which must be disrupted before interaction with U6 can occur (Madhani and Guthrie, 1992; Fortner *et al.*, 1994).

1.10.4 Interaction of U6 snRNA with the Substrate.

Adjacent to the region forming helix I in U6 snRNA is the conserved sequence ACAGAG, which is essential for splicing. Mutation of intron position 5, which activates aberrant cleavage sites (Fouser and Friesen, 1986) can be suppressed by restoring complementarity to the second position of this element (C in ACAGAG; Kandels-Lewis and Séraphin, 1993; Lesser and Guthrie, 1993). This is consistent with

the cross-linking data (Sontheimer and Steitz, 1992). The suppression restores cleavage to the normal 5' splice site. In contrast when the intron position 5 mutant is combined with a mutant U1 snRNA, also restoring complementarity, splicing at the aberrant site still occurs (Séraphin *et al.*, 1988; Séraphin and Rosbach, 1990; Silicano and Guthrie, 1988). Thus U1 and U6 base-pair with the same region of the substrate. It is unlikely that these interactions occur simultaneously; the most likely scenario is that U1 is displaced after tri-snRNP entry into the spliceosome. Indeed, disruption of the U1/5'splice site helix is required for spliceosome assembly (Konforti *et al.*, 1993). These data have led to the hypothesis that the U1 interaction determines only the general region of the substrate which contributes the 5' splice site, and that later base-pairing interactions with U6 (and possibly U5) determine precisely which phosphate is attacked in the first transesterification.

1.10.5 Tertiary Interactions.

U2 snRNA bulges out of U2/U6 helix I. Madhani and Guthrie (1994b) mutagenised both U2 snRNA in the region of this bulge and the ACAGAG of U6, and simultaneously selected for mutants in one and suppressors in the other (a technique called randomisation/selection). They showed that the last position in ACAGAG and the U2 bulge exhibited covariance, suggesting a non-Watson-Crick molecular interaction. This would bring together two snRNA elements essential for step II, and together with the U6-intron crosslinking data suggests that these snRNA regions are part of the active site in step II. A second tertiary interaction has been detected in yeast by reciprocal suppression between the first and last intron positions of the intron, which are invariantly G (Parker and Silicano, 1993; Chanfreu *et al.*, 1994).

1.10.6 Catalysis of RNA Conformational Change: The DEAD-box Putative RNA Helicases.

The dynamic RNA-RNA interactions which occur in the spliceosome are modulated by protein:RNA interactions. Five splicing factors identified in yeast (Prp2p, Prp5p, Prp16p Prp22p, and Prp28p) bear the primary sequence motif characteristic of the DEAD/H-box protein family (Fuller-Pace and Lane, 1992). Of the 30 or more known members of this family, only four including the prototype eIF-4A have been

demonstrated to have RNA unwinding (helicase) activity (Ray *et al.*, 1985; Rozen *et al.*, 1989; Hirling *et al.*, 1989; Lain *et al.*, 1990; Lee and Hurwitz, 1993). In other members of the family including the five splicing factors, helicase activity has proven difficult to demonstrate. However, they are predicted to have analogous functions, driving conformational change of RNA:RNA duplex.

Prp5p and Prp28p have been described already, and have been shown genetically to be involved in U2 snRNP association with the commitment complex, and U4/U6 dissociation, respectively. The Prp2p, Prp16p and Prp22p are members of the DEAH subfamily, and bear significant similarity to one another over approximately 450 residues. None has demonstrated helicase activity, but Prp2p and Prp16p bind RNA and have RNA-stimulated ATPase activity *in vitro* (Kim *et al.*, 1992; Schwer and Guthrie, 1991). These two proteins have been shown to bind to the spliceosome at different stages of the splicing pathway.

In a heat-inactivated extract from a *prp2* strain a fully assembled spliceosome (complex I) is able to form, but the sequence of events is blocked prior to the conformational change to the active complex II (Lin *et al.*, 1987). Given a supply of ATP, the cycle of Prp2p binding, ATP hydrolysis and dissociation is very rapid (King and Beggs, 1990; Plumpton *et al.*, 1994), however the spliceosome/Prp2p complex can be stalled by two means. A spiceosome assembled in the absence of Prp2p will bind to purified Prp2p in the absence of ATP (Kim and Lin, 1993; Teigelkamp *et al.*, 1994). Step I does not proceed and the spliceosome is stalled until exogenous ATP is added when pre-mRNA is chased to products. Stalled spliceosome/Prp2p complexes also are the result of the mutation to *PRP2* which causes a serine (S) to leucine (L) amino acid change in the conserved SAT motif. In this case ATP does not enable splicing to proceed, and the stalled complex is stable *in vitro* (Plumpton *et al.*, 1994). The phenotype of this mutation is dominant negative: the mutant inhibits splicing and is toxic when overproduced in wild-type yeast cells due to competition with wild-type Prp2p (Plumpton *et al.*, 1994). Other mutations in the motifs conserved amongst DEAD-box family members cause this same phenotype (Plumpton *et al.*, 1994; Flinn, 1994). In the stalled spliceosomes, Prp2p can be crosslinked to the substrate pre-mRNA, suggesting that this is one of the RNAs towards which the putative helicase activity of Prp2p is directed (Teigelkamp *et al.*, 1994).

Extracts depleted of Prp16p accumulate lariat intermediates, so there is no absolute requirement for Prp16p until after step I (Schwer and Guthrie, 1991). The *in vitro* activities of Prp16p protein are very similar to Prp2p: Prp16p associates with spliceosomes in the absence of ATP, but ATP is required for the completion of splicing (Schwer and Guthrie, 1991). The *PRP16* gene was originally isolated as the mutant *prp16-1* (Couto *et al.*, 1987). This allele suppresses a branchpoint mutation (C259) in an intron-*HIS4* gene fusion which usually inhibits both steps of splicing, preventing complementation of a his$^-$ phenotype. *prp16-1* allows step II to occur in spite of the branch site mutation. The defective protein has wild-type affinity for the spliceosome but greatly reduced ATPase activity, and the allele is partially dominant negative and partially functional as Prp16p (Schwer and Guthrie, 1992). Other mutations in *PRP16* which also suppress BPS mutations also have reduced ATPase activity (Burgess and Guthrie, 1993), suggesting that any decrease in ATP hydrolysis can mediate suppression. These results have led Guthrie and co-workers to propose a proofreading mechanism, whereby ATP hydrolysis activates a discard pathway for aberrant lariat intermediates which results in their destruction. The decreased rate of ATP hydrolysis is envisaged to allow aberrant intermediates more time to undergo step II. Consistent with this the lariat intermediate is turned over rapidly in wild-type cells but is relatively stable in the *prp16-1* mutant (Burgess and Guthrie, 1993). Whether Prp2p has an analogous function in step I has yet to be determined.

1.10.7 The Active Sites.

The end result of these conformational changes is the assembly of a catalytic active site. Before considering whether step I and step II occur in the same or different active sites, it is relevant to introduce a second category of self-splicing intron, the group I introns. As with group II, these excise by folding into autocatalytic tertiary structures, but excision follows a slightly different mechanism. The group which attacks the group I intron 5' splice site is the 2' hydroxyl of a non-covalently bound guanosine nucleotide. The second step is similar to the group II/spliceosome mechanism. The two steps of group I splicing have opposite stereochemical requirements: the first step requires R_p chirality, and the second step requires S_p chirality. This has been deduced using phosphorothioate analogues (McSwiggen and

Cech, 1989; Suh and Waring, 1992). From this, and from the fact that the attacking group in the first step and the leaving group in the second step are both guanosine 2' hydroxyl groups, it seems likely that the two reactions occur in a single reactive site and that the second step is simply the reverse of the first, after exchange at a G-binding pocket of the guanosine nucleotide for the G at the 3' splice site. In contrast, the spliceosome reaction requires the same chirality at each step (Moore and Sharp, 1993), and the attacking/leaving groups in each step are different. Therefore, unlike group I splicing, in the spliceosome steps I and II cannot occur in one active site. Step II cannot be simply the reverse of step I, as this would contravene the principle of microscopic reversibility. The reactions in the spliceosome probably take place in (a minimum of) two active sites; these may be overlapping.

After completion of step II, spliced mRNA is released from the spliceosome and exported from the nucleus. Spliceosome disassembly is an active, ATP-requiring process and the lariat intron product is debranched by a specific enzyme activity, and degraded. It is generally assumed that *in vivo* snRNPs and other spliceosomal components are recycled.

1.11 *PRP8.*

1.11.1 The Identification of *PRP8.*

The Prp8p splicing factor of yeast is the subject of this thesis. Prp8p is the product of the *PRP8* gene, which was isolated by Hartwell and originally known as *RNA8* after a complementation group of mutants which had the Rna$^-$ phenotype (Hartwell, 1967; Hartwell *et al.*, 1970). These mutants were defective in stable rRNA accumulation, however synthesis of rRNA precursors was unaffected, leading to the notion that cells had an rRNA processing defect (Shulman and Warner, 1978). It is believed that defective pre-mRNA splicing manifests this phenotype due to the predominance of introns in ribosomal protein genes. Inability to excise these introns prevents production of ribosomal proteins. Ribosomal RNA is not properly processed because the substrates of rRNA processing are complexes of rRNA with ribosomal proteins, leading to the rapid turnover of rRNA (Rymond and Rosbash, 1992). The

RNA genes were renamed *PRP* when their products were shown to be required for pre-mRNA splicing (Lustig *et al.*, 1986).

Prp8 protein is a 280 kD integral component of the U5 snRNP (Lossky *et al.*, 1987; Whittaker *et al.*, 1990). It has been shown in immunoprecipitation studies and by affinity purification of spliceosomes to be a component of the U4/U6.U5 tri-snRNP and of assembled spliceosomes both before and after the transesterification reactions (Lossky *et al.*, 1990; Whittaker *et al.*, 1990; Teigelkamp *et al.*, 1995a). Prp8p is required for the association of the U5 and U4/U6 snRNPs and of the tri-snRNP with the spliceosome (Brown and Beggs, 1992). For several years after *PRP8* was cloned and sequenced (Jackson *et al.*, 1988; Hodges *et al.*, 1995), no homology could be detected to any other protein in the database. However, antibodies raised against regions of Prp8p bound to proteins of similar large size (more than 200 kD) extracted from human (Anderson *et al.*, 1989; Pinto and Steitz, 1989; Garcia-Blanco *et al.*, 1990), rat (Guialis *et al.*, 1991), mouse (Hodges *et al.*, 1995), *Drosophila melanogaster* (Paterson *et al.*, 1991), *Caenorhabditis elegans* (Hodges *et al.*, 1995), tobacco and pea (Kulesza *et al.*, 1993).

A *PRP8* homologue in *Ceanorhabditis elegans* was sequenced as part of a cosmid in the Nematode Sequencing Project. The *C. elegans* nucleotide sequence is 60% identical to yeast *PRP8*, and the deduced proteins contain 61% identical and 77% similar amino acids (Hodges *et al.*, 1995). The region of similarity spans virtually the entire yeast *PRP8* gene, and the predicted proteins are co-terminal at their C-termini. A *PRP8* homologue from maize has been cloned and sequenced (P.E. Hodges, J.W.S. Brown, and J.D. Beggs, unpublished), and in the database are several expressed sequence tags from rice, human, *Arabidopsis thaliana* and *Plasmodium falciparum*. All show a high degree of homology to the yeast gene. Such strict conservation is unique amongst known splicing factors (Hodges *et al.*, 1995) and of an order comparable to such important genes as *CDC28* (Shea *et al.*, 1994).

1.11.2 The Role of *PRP8*.

Metabolic depletion of Prp8p causes levels of U4, U5 and U6 snRNAs to decline dramatically, causing a lethal splicing defect (Brown and Beggs, 1992). A role for Prp8p in stabilizing snRNAs in the tri-snRNP is also suggested by genetic data. The

prp8-9 allele was isolated as a suppressor of the cold-sensitive *prp28-1* mutant. The Prp28p putative RNA helicase is proposed to unwind the U4/U6 duplex; Prp8p may counterbalance this activity (Strauss and Guthrie, 1991). The *prp8-1* allele has been suggested to cause degradation of U4, U5 and U6 snRNAs similar to depletion of the protein (Brown and Beggs, 1992). *prp8-1* is suppressed by mutation in another putative helicase, *DED1* (Jamieson *et al.*, 1991).

Whittaker and Beggs (1991) demonstrated a physical interaction between Prp8p and pre-mRNA by UV crosslinking. This interaction requires assembly of the substrate into a spliceosome, but is not dependent on the 3' splice site. Analysis of the crosslinked RNAs by RNase T1 treatment suggested that both the 5' splice site and intron 3' regions were potential sites of interaction with Prp8p (Teigelkamp *et al.*, 1995a). The exact positions of Prp8p interaction were mapped in substrates with site-specifically incorporated 4-thiouridine nucleotides (Moore and Sharp, 1992). Prp8p is apparently in contact with quite long regions of the substrate, upstream of the 5' splice site (at least 8 nucleotides) and around the 3' splice site (at least 13 nucleotides; Teigelkamp *et al.*, 1995b). The interaction with the upstream exon is established before step I, and with the downstream exon (and also the branchpoint) after step I. Thus the contacts made by Prp8p and by U5 snRNA with the substrate are in accordance, both temporally and spatially. In common with U5, the Prp8p interaction with exon regions is not sequence specific (Teigelkamp *et al.*, 1995b). The mammalian Prp8p homologue p220 is also a U5 snRNP component (Anderson *et al.*, 1989; Pinto and Steitz, 1989), and can be crosslinked to a uniformly-labelled substrate (Garcia-Blanco *et al.*, 1990) or specifically to the upstream exon position -2 (Wyatt *et al.*, 1992) or the BPS (MacMillan *et al.*, 1994).

Genetic approaches have also highlighted functional interactions of Prp8p both with U5 snRNA and with the 3' splice site. U5 loop I mutations which impede growth are synthetic lethal in combination with conditional *prp8* alleles (Guthrie and Patterson, 1988; Frank *et al.*, 1992). Other *prp8* alleles affect the preference for uridine-rich 3' splice sites in yeast introns. The *prp8-101* mutation was discovered in a genetic screen which involved competition between two alternative 3' splice sites. The proximal 3' site (closest to the 5' splice site) was uridine-rich and used preferentially over a distal, adenosine-rich 3' site. This intron was fused to the *CUP1* gene; splicing to the distal,

disfavoured 3' splice site was required for expression of *CUP1* and copper resistance (Lesser and Guthrie, 1993). *prp8-101* enables CUP1 expression, and also causes defective splicing of any yeast intron which requires polypyrimidine tract recognition, such as the *TUB3* gene (Umen and Guthrie, 1995).

The experiments described above identify functions of the Prp8p/p220 splicing factor at multiple stages of the splicing process. Prp8p functions prior to spliceosome assembly to stabilise snRNAs, during assembly of the tri-snRNP and the spliceosome, during the U4/U6 destabilisation events, and during both step I and step II of splicing. The strict conservation of this factor is a reflection of its multifunctionality. The consensus of this work is that Prp8p acts as a molecular scaffold, conferring stability and support to multiple RNA:RNA interactions in spliceosomal complexes (Whittaker and Beggs, 1991). One of these RNA:RNA interactions is likely to be that of the U5 snRNA with the substrate, which is intrinsically weak. An activity to bind the free upstream exon between step I and step II and actively retain it in the spliceosome has long been sought, and the Prp8p/U5 snRNA pair is in the right place at the right time (Teigelkamp *et al.*, 1995b). Conceivably, the Prp8p/U5 snRNA pair tethers the upstream exon before step I, and juxtaposes it at the 3' splice site ready for step II. Prp8p contains none of the sequence motifs implicated in RNA binding, and does not appear to exhibit non-specific RNA binding activity (Whittaker and Beggs, 1991; Teigelkamp *et al.*, 1995a; Teigelkamp *et al.*, 1995b), suggesting that an RNA-binding conformation only forms within the spliceosome, or that complementarity is specific to RNA structures which only form in splicing complexes.

1.11.3 The N-terminal Region of *PRP8*.

Following the methionine start codon of Prp8p is a stretch of 85 residues (designated henceforth the N-terminal domain) containing 16 acidic amino acids (aspartate or glutamate) and no basic amino acids. Within this region are four copies of a proline-rich motif conforming to the consensus LP_nG (where n=5 to 8). In one copy of the repeat the last amino acid of this motif is S (serine) instead of G (glycine). Directly adjacent to this proline-rich domain is a region (residues 86 to 141) in which basic residues (lysine, arginine, or histidine) predominate amongst the charged amino acids (23 out of 31 charged amino acids).

Antibodies raised against a synthetic peptide copy of amino acids 2 to 35 of Prp8p (containing two copies of the LP_nG motif) fail to detect Prp8p homologues in other species, although they react with yeast Prp8p strongly. This is consistent with the predicted sequence of the Prp8p homologue of *C. elegans*, which is co-terminal with yeast Prp8p at the C-terminus, and closely homologous through the bulk of the protein including the basic domain, but which contains no equivalent of the proline-rich acidic N-terminal domain. All start codons in *C. elegans* are preceded by a 3' acceptor splice site, because maturation of pre-mRNA involves *trans*-splicing of short leader sequences onto the 5' end of each mRNA (Blumenthal, 1995); this is true also of the Prp8p homologue. The 3' splice sites for *cis*- and *trans*-splicing are the same, so the formal possibility that extra peptide sequence is added at the N-terminus by splicing events cannot be discounted. However, there is no exon with the potential to encode a proline-rich peptide with homology to the N-terminal region of yeast Prp8p in 6.2kb of sequence preceding the *C. elegans* coding region (Hodges *et al.*, 1995). The maize genomic *PRP8* clone (Hodges *et al.*, 1995) encodes a predicted protein 84% identical to the *C. elegans* protein, and again similarity to yeast Prp8p does not extend into the N-terminal domain. It cannot yet be ascertained whether the genes isolated from maize or *C. elegans* encode functional proteins; in maize there are definitely multiple copies of the gene (J.W.S. Brown, personal communication). So although it cannot yet be said with certainty that homologues of Prp8p lack a proline-rich N-terminal domain, there is preliminary indication that this region is unique to yeast Prp8p.

1.12 Proline-rich Regions in Other Proteins.

Proline-rich regions are commonly found in both prokaryotic and eukaryotic proteins, and frequently occur as multiple tandem repeats (for review see Williamson, 1994). The unusual nature of proline, with its 3-carbon side chain cyclised back onto the amide position, strongly influences the properties of proline-rich regions. The peptide backbone is unable to rotate about the C^{α}-N of proline. Also the N-C^{δ} group places steric hindrance on the amino acid preceding proline, further reducing the rotational freedom of a proline-rich peptide. Proline lacks an amide proton and thus cannot act as a hydrogen bond donor. For these reasons, proline does not readily form

α-helical or β-sheet secondary structures. Where proline does occur in an α-helix, it creates a kink in that structure, which in the case of transmembrane helices has been proposed to have functional significance (Williams and Deber, 1991). In solution, homopolymers of four or more prolines adopt an extended helical conformation, with three amino acids per turn. This is the polyproline II helix, characteristic of collagen (Cowan and McGavin, 1955). This conformation is regularly observed in crystal structures, typically in a short peptide containing one or more prolines at the protein surface. Presumably, proline-rich regions (which do not readily crystallise) in general form extended peptide arms outside globular centres of a protein.

The survey of Williamson (1994) concluded that many proline-rich regions function to bind other proteins, and particularly to form non-stoichiometric, non-specific interactions. Binding partners are often proline-rich too. One example is RNA polymerase II which contains 26 or 27 copies of a motif of consensus YSPTSPS at its C-terminus. This domain has been suggested to interact with other components of the transcription initiation complex and retain them in a complex with indefinite stoichiometry but a limited spatial range (Sigler, 1988). Such elements might include the activation domains of transcription factors which are often proline-rich (Mermod *et al.*, 1989; Gerber *et al.*, 1994).

In the transcriptional initiation complex, or the spliceosome, or any other multicomponent complex where a protein needs to interact with numerous different targets, non-specific, non-stoichiometric binding might be advantageous. Proline in a peptide ligand dictates that it forms a constrained, extended structure, thus binding is faster than with two complementary protein surfaces. A proline-rich peptide has less entropy and therefore a lower entropic barrier to binding, favouring the overall binding energy over a similar peptide lacking prolines. Also, the enthalpy of binding can be advantageous because proline (as acceptor) forms a stronger hydrogen bond than other amino acids because its carbonyl is more electron-dense.

Although binding events involving proline-rich regions have been suggested to be non-specific, one example of specific binding is familiar, that of proline-rich peptide ligands to the SH3 domain of the mammalian nucleotide-release factor, Sos. This interaction stringently requires prolines at positions 2, 7, and 10 of an 11-amino acid consensus (Ren *et al.*, 1993). Specificity is conferred by positions 1 and 11; presumably

the prolines function to impart a binding-proficient conformation on the motif (Rozakis-Adcock *et al.*, 1993). The solution structure of an SH3 domain (in this case from phosphatidylinositol 3-kinase) and its bound proline-rich peptide ligand has been solved, and confirms that the bound peptide is in the polyproline II helix conformation (Yu and Schreiber, 1994).

1.12.1 Other Proline-rich Splicing Factors.

Prp8p is not the only splicing factor which contains a repetitive proline-rich element: the C-terminal third of the mammalian U2 snRNP protein SAP62 contains 22 repeats of the sequence GVHPPAP, and adjacent domains of this protein are also proline rich (Bennett and Reed, 1993). Prp11p, the yeast homologue of SAP62 lacks the proline-rich motifs, and homology between the two spans only the N-terminal two-thirds of the mammalian protein (Chang *et al.*, 1988; Bennett and Reed, 1993).

The third proline-rich splicing factor is mammalian PSF, which was identified by its association with polypyrimidine tract-binding factor (PTB). PTB was first characterised as a factor which could be UV cross-linked to the polypyrimidine tract of pre-mRNAs. Furthermore, with a variety of substrates the extent of interaction with PTB correlated with the efficiency of prespliceosome complex formation, and overall splicing efficiency, suggesting that PTB was a splicing factor (Garcia-Blanco *et al.*, 1989). Passage of a HeLa nuclear extract over a poly(U)-Sepharose matrix depletes those factors which bind the polypyrimidine tract. Two factors are required to restore splicing activity to such a depleted extract by complementation; they are U2AF, and a heterodimer of PTB and second protein termed PTB-associated factor (PSF; Patton *et al.*, 1991; Patton *et al.*, 1993; Gil *et al.*, 1991). PSF is identical to SAP 102 (Bennet *et al.*, 1992b). PTB and PSF interact directly *in vitro*, and PSF has independent binding affinity for pyrimidine-rich sequence, and is an essential splicing factor (Patton *et al.*, 1993). Several different cDNAs encoding PSF have been isolated, suggestive of alternative splicing and polyadenylation, but all isoforms have two copies of the 80-amino acid RNP RNA recognition motif, and an N-terminus which is rich in proline and glutamine, but which is not repetitive (Patton *et al.*, 1993). This region also contains three RGG motifs, which may mediate RNA recognition (Kiledjian and Dreyfuss, 1992). The function of the proline-rich region in PSF is unknown.

The preferred binding partner of PTB is not a typical polypyrimidine tract (in contrast to U2AF) but a special type found in alternatively spliced 3' splice sites subject to negative regulation. PTB bound to these sites represses their selection in splicing (Singh *et al.*, 1995). This type of polypyrimidine tract does not bind U2AF with high affinity. Consistent with a role as a regulator of alternative splicing, PTB is itself alternatively spliced, is expressed in a tissue- and developmental stage-specific fashion, and is probably not essential for splicing. The existence of a heterodimer in mammalian cells, with both components (PTB and PSF) having independent affinity for the polypyrimidine tract, one being an essential splicing factor and one a repressor of splicing, is remarkable.

1.13 This Thesis.

The above sections have introduced the Prp8p pre-mRNA splicing factor of budding yeast. The aim of this thesis was to analyse the role of Prp8p, and to identify its functional domains. To do this, truncation mutants were created which retained certain functions of wild-type Prp8p. In a wide variety of cellular systems, partial loss of function is often the result of truncation mutation, and occasionally the highly informative dominant negative phenotype (Herskowitz, 1987). These studies have concentrated on the proline-rich, acidic region at the N-terminus of Prp8p, which on the basis of its physical properties is predicted to form a structural domain. Many of the proline-rich regions of other proteins whose function is known act as protein binding sites (Williamson, 1994).

In the course of these studies, our research group was made aware of *dbf3-1*, a mutant allele of *PRP8* with a novel phenotype. Unlike other *prp8* alleles, *dbf3-1* affects cell cycle progression (Johnston and Thomas, 1982a and 1982b; Shea *et al.*, 1994). This phenotype implied novel functions for Prp8p (Shea *et al.*, 1994) additional to those previously characterised in pre-mRNA splicing; this possibility was investigated.

In this thesis I present:

(i) The demonstration that deletion of part of the repetitive proline-rich Prp8p N-terminus causes partial loss of function, and that the defect caused by deletion can be overcome by overexpression.

(ii) The characterisation of a yeast mutant viable in the absence of the Prp8p N-terminus. The snRNP profile of this strain is suggestive of a role for the proline-rich region in promoting splicing complex assembly.

(iii) A preliminary characterisation of the minimum sequence which can functionally replace the wild-type N-terminus. Proline, rather than acidity, seems to be the most important feature of the N-terminus.

(iv) The demonstration that *prp8* alleles including *dbf3-1* and others with a cell cycle defect all affect pre-mRNA splicing.

(v) The allele-specific suppression of *dbf3-1* by a cDNA copy of the *TUB1* gene. This suppression separates the cell cycle and splicing phenotypes, and suggests a possible basis of the effect of *dbf3-1* (and by analogy the other cell cycle *prp8* mutants) on cell cycle progression.

This research supports a model in which the apparent cell cycle phenotype of *dbf3-1* is a consequence of a defect in pre-mRNA splicing. To come to this conclusion, some background of the yeast cell cycle is obviously helpful, and the following sections are included to provide this. These sections are not intended to be a fully comprehensive review, but rather an overview of cell cycle research in yeast. They are extracted from several recent reviews (Nasmyth, 1993; Reed, 1992; Pines, 1995 and 1994; Koch and Nasmyth, 1994; Tye, 1994). References to experimental observations, where not included in one of these reviews, are given.

1.14 An Overview of the Yeast Cell Cycle.

The proliferation of a budding yeast cell requires two key events: the high-fidelity replication of its genetic complement (S-phase) followed by segregation of one of the copies of the genome to a daughter cell (M-phase). Obviously, these events and other associated processes (for example bud formation, construction of a mitotic spindle, and the achievement of a certain cell size) must occur in a specific order, and it is unsurprising that yeast (like all other eukaryotes) has an elaborate system to regulate this cell cycle. The co-ordinator and master regulator of the cell cycle is the product of the *CDC28* gene, a cyclin-dependent kinase (CDK) whose activity is regulated by association with cyclin subunits at specific times during the cell cycle. Different cyclins are present at different phases of the cell cycle because cyclins in general are unstable, and subject to targeted degradation. Cyclins impose substrate specificity on the Cdc28 kinase, whose activity drives the cell through successive cell cycle checkpoints. Nine different cyclins are thought to sequentially interact with Cdc28p to effect cell cycle regulation; they are the Cln1 to 3 and Clb1 to 6.

1.14.1 The Cyclin Cascade.

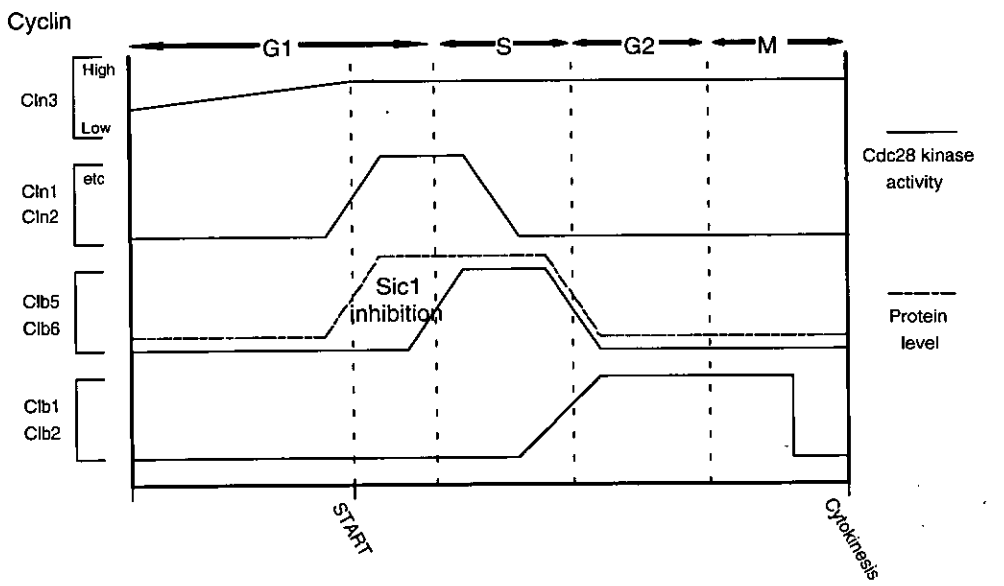


Figure 1.3: The Cyclin Cascade.

Under non-adverse conditions, a yeast cell undergoes a decision process during G_1 (the phase between the end of M-phase and the beginning of S-phase) which commits it to undergoing subsequent rounds of S-phase and M-phase. Prior to this decision, multiple cell fates are possible; these include a quiescent state in which cells are relatively heat-tolerant, or for haploid cells response to mating pheromone, or for diploid cells the initiation of meiosis. After the decision to proceed to S-phase, mitosis is the only fate of the cell, even in the absence of nutrients. The point at which this decision is made is designated START, and results in transcriptional changes and the recruitment of a large group of proteins to DNA replication. Also at START, the spindle pole bodies replicate, the process which leads to bud emergence is initiated, and the responsiveness to mating pheromone ceases. At START, Cdc28p is associated with a set of G_1-specific cyclins, including Cln1p and Cln2p. These two cyclins are constitutively unstable, and accumulate during G_1 (as cells reach a certain size) due to transcriptional derepression. During S-phase Clb5p- and Clb6p-Cdc28 kinase complexes become active. Clb5p and Clb6p accumulate during G_1 in parallel with the Cln cyclins, but until the start of S-phase Clb5p- and Clb6p-Cdc28 kinase complexes are associated with the specific CDK inhibitor (CDI), Sic1p. As Cln complexes become active, Sic1p becomes conjugated to ubiquitin and degraded in the proteasome, causing activation of Clb5p and Clb6p activity. During G_2 (the phase between the end of S-phase and the beginning of M-phase) a third set of CDK-cyclin complexes become active, containing Clb1p and Clb2p. As Clb1/2 complexes become active, *CLN1* and *CLN2* transcription is repressed; it is through feedback loops such as these that the cell is able to switch between different cyclin/CDK complexes. Clb1p and Clb2p are only transcribed after START, and do not persist after M phase, again because of ubiquitin-targeted proteolysis. This proteolytic activity begins in M-phase, persists through G_1 and is inactivated at START in response to Cln1/2-Cdc28p activity. As Clb cyclins are degraded in M-phase, so *SIC1* expression commences ready for the subsequent G_1.

It can be seen that cell cycle regulation involves a cascade of cyclins (Figure 1.3), each causing the activation of the next, and leading to the production of two cells (mother and daughter) by mitosis. The commitment events during G_1 which initiate the cell cycle are still somewhat unclear, but are thought to depend on Cln3p. This cyclin is unique in that its mRNA and protein levels do not vary widely during the cell cycle,

although during G_1 rates of *CLN3* transcription are proportional to cell size, as is typical of any aperiodically expressed gene. Cln3p is apparently able to monitor cell size during G_1 (Tyers *et al.*, 1993), and may be at the top of the cyclin cascade, activating *CLN1* transcription when the cell reaches a certain size. Possibly, once the gradually increasing rate of *CLN3* transcription exceeds a certain threshold, Cln3-Cdc28 kinase activity becomes sufficient to initiate the cyclin cascade.

1.14.2 Transcription and the Cell Cycle.

Underlying the cyclin cascade are transcriptional changes involving several different transcription factors. As many as 250 transcripts may be regulated in the yeast cell cycle. Genes transcribed exclusively in late G_1/early S phase can be subdivided into two groups according to the *cis*-acting sequences found in their promoters. The SCB element (for Swi4/Swi6 cell cycle box, CACGAAA) is found upstream of several genes including *CLN1* and *CLN2*, and the MCB element (for *Mlu*I cell cycle box, ACGCGTNA) is found upstream of *CLB5* and *CLB6*, and others including many genes involved in DNA synthesis. Both elements can confer cell cycle phase-specific expression on otherwise inactive promoters.

These elements are bound by the SBF and MBF transcription factors, which are heterodimeric complexes between homologous DNA-binding subunits (Mbp1p in MBF and Swi4p in SBF) and a common regulatory subunit, Swi6p. Once Cln1/2 cyclins become active, MBF and SBF become part of a positive feedback loop which acts to maintain Cln1/2-Cdc28 activity: MBF is activated by Cln1/2p, and SBF and MBF induce *CLN1/2* transcription. Later, in G_2 SBF- and MBF-regulated genes are repressed. In the case of SCB-containing promoters this repression is in response to Clb1/2-Cdc28p activity; with MCB genes this repression is apparently independent of cyclins Clb1/2. The expression of cyclin genes in response to these transcription factors in late G1/early S phase is required for normal cell cycle progression. The many other genes which are subject to the same regulation such as those involved in DNA replication are presumably expressed at this time for reasons of economy, although as their gene products are in general stable, cells would probably still be viable if expression of these genes was constitutive. Even cyclin genes can be placed under constitutive transcriptional regulation without causing lethality.

It seems likely from these observations that MBF and SBF interact with cyclin/CDK complexes. For example, as cells obtain the requisite size for cell cycle commitment, Cln3/CDK presumably activates SBF to initiate *CLN1/2* transcription. The molecular basis of this putative interaction is not known. It may be that subunits of SBF and MBF are directly phosphorylated by CDK. Alternatively, the cyclin/CDK complex may associate with the MBF- or SBF-containing transcription complex and phosphorylate a third party, such as the C-terminal domain (CTD) of RNA polymerase, which requires to be phosphoryalated before transcriptional initiation.

Less is known about transcription events later in the cell cycle, but the generality that cyclin/CDK complexes activate particular transcription factors which stimulate expression of appropriate genes, in particular cyclin genes, seems to hold. Histone genes are expressed from the onset of S-phase, and *CLB3* and *CLB4* slightly later (Clb3 and Clb4 cyclins are not included in figure 1; they accumulate just in advance of Clb1 and Clb2 cyclin, and it is not yet clear whether they possess functions distinct from the other G_2 cyclins). A set of genes required for M-phase onset including *CLB1* and *CLB2*, and also other genes whose products act at the end of M-phase and in early G_1 such as *CDC46* and *SWI5*, are expressed from the beginning of G_2, at the time when Clb1/2-CDK activity commences. Both the SWI5 and CLB2 promoters contain binding sites for the heterodimeric transcription factor SFF/Mcm1p, which is thus a likely candidate for a coordinate regulator of G_2-specific transcription. Swi5p is itself a transcription factor; along with the closely related Ace2p it is responsible for a late M/early S phase program of transcription. Genes under this control include chitinase (*CTS1*) and other genes involved in cell separation, the *HO* recombinase, and possibly *SIC1*. Swi5p is synthesised in G2 but not active until M-phase because its nuclear localisation signal (NLS) is masked by CDK-directed phosphorylation. As Clb cyclins are destroyed in M-phase, the NLS is dephosphorylated and Swi5p enters the nucleus.

1.14.3 Replication Licensing Factors.

Initiation of DNA replication must occur only once between each M-phase. The cell guarantees that this is the case by means of replication licensing factors: factors which are absolutely required for DNA replication, that are inactivated or destroyed during S-phase, and that gain access to the nucleus during M-phase (Tye, 1994). Thus

DNA replication is under *trans*-acting positive control. The first indication of the existence of such factors came from the work of Laskey and co-workers in the *Xenopus* system. Incubation of sperm chromatin with *Xenopus* egg extract results in assembly of a nuclear envelope and a single round of DNA replication. Re-replication does not occur unless M-phase takes place, or the nuclear membrane is made permeable with detergents, suggesting the existence of a diffusible factor in the egg extract. More recent studies have demonstrated that Xenopus licensing factors are homologues of a group of proteins originally discovered in yeast (the Mcm proteins; Chong *et al.*, 1995; Madine *et al.*, 1995; Kubota *et al.*, 1995).

CDC46/MCM5 is one of a family of genetically interacting genes (the others are *CDC45*, *CDC54* and *CDC47*) identified as mutants which block the G_1/S phase transition. Cdc46p is of particular interest because its pattern of expression and localisation within the cell are striking: the gene is expressed in G_2 and the protein appears in the nucleus at M-phase, where it persists until just after the G_1/S boundary. The two other members of the MCM family, *MCM2* and *MCM3*, have structural similarities to each other and to *CDC46*, including a shared ATPase domain similar to that of DNA helicases. In addition *MCM2* has a zinc-finger motif, and *MCM3* has an archetypal NLS. These genes were identified as mutant alleles with the mcm (for minichromosome maintenance) phenotype; the destabilisation of plasmids containing the ARS element. Conditional *mcm2* and *mcm3* mutants block cell cycle progression late in S-phase, and cause reduced rates of initiation at particular ARS elements.

All three MCM genes are essential for viability and have the same pattern of periodic localisation. Inside the nucleus these proteins are very tightly associated with DNA. The periodic localisation of MCM proteins divides the cell cycle into a period where DNA replication is possible, and a period where it is not. An attractive model is that Mcm proteins respond to the activity of cyclin/CDK, and act to unwind origins of replication, a process required for the initiation of DNA replication.

CHAPTER 2.
MATERIALS AND METHODS.

2.1 MATERIALS.

2.1.1 Suppliers of Laboratory Reagents.

Standard laboratory reagents: Sigma, Fisons, BDH, Bio-Rad, Serva.

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

Radiochemicals: Amersham International.

Deoxyribonucleotides and ribonucleotides: Pharmacia and Sigma.

Acrylamide and *N,N'*-methylene bisacrylamide: BDH chemicals, Electran grade.

Urea: Fisons.

Agarose: Boehringer Mannheim (Multi Purpose), Gibco BRL (Low Melting Point Agarose).

Phenol: Rathburn Chemicals (glass-distilled, water-saturated).

Media Reagents: Difco Laboratories, Sigma.

2.1.2 Growth Media.

Except where otherwise indicated, solutions were autoclaved and stored at room temperature. Quantities are for 1 litre volume. Antibiotics and amino acids were stored as filter sterilized, concentrated stocks, and were added to media immediately prior to use. To make agar plates, bacto-agar was added to 2% (w/v) prior to autoclaving.

2.1.2.1 *E.coli* Media.

Luria-Bertani medium (LB): 10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl, adjusted to pH 7.2 with NaOH.

LB-amp: LB containing 100µg/ml ampicillin.

M9 medium: 6g Na_2HPO_4, 3g KH_2PO_4, 0.5g NaCl, 1g NH_4Cl, 0.25g $MgSO_4.7H_2O$, 2g glucose. Amino acids were added after autoclaving to 20µg/ml.

2X YT-amp: 10g bacto-peptone, 10g bacto-yeast extract, 5g NaCl, 100µg/ml ampicillin.

2.1.2.2 Yeast Media.

YPDA: 10g bacto-peptone, 10g bacto-yeast extract, 20g glucose, 20mg adenine sulphate.

YPGRA: as YPDA, except 20g galactose and 20g raffinose instead of glucose.

YMGlu: 6.7g yeast nitrogen base without amino acids, 20g glucose, supplemented with amino acids as required to 20µg/ml.

YMGR: as YMGlu except 20g galactose and 20g raffinose instead of glucose.

YMGlucas/YMGRcas: YMGlu or YMGR plus 10g vitamin assay casamino acids.

Sporulation: 10g potassium acetate, 1g bacto-yeast extract, 0.5g glucose, supplemented with amino acids as required to 2µg/ml.

YMGlyLac: 6.7g yeast nitrogen base without amino acids and 10g vitamin assay casamino acids were dissolved in 700ml dH_2O and autoclaved. When cool, 100ml sterile 20% (v/v) lactate (adjusted to pH5.7 with KOH at 0°C) and 100ml sterile 20% (v/v) glycerol were added. Volume was made up to 1l with either 100ml sterile dH_2O, or 100ml sterile 20% (w/v) galactose (for induction of *GAL1* promoter). Lactic acid was obtained from Sigma (Cat. No. L-1250).

2.1.2.3 Use of Microtubule-directed Drugs.

Nocodazole was purchased from Sigma; methyl benzimidazol-2-yl carbamate (MBC) was the generous gift of Colin Gordon, MRC Human Genetics Unit, Edinburgh. To prepare plates containing either nocodazole or MBC, media were autoclaved and allowed to cool to 55°C. DMSO, and/or a 10mg/ml stock solution of the drug in DMSO were added while stirring continuously, and plates poured. Final concentration of DMSO in all plates was 1%.

2.1.3 Bacterial Strains.

Strain NM522 (Δ*(lac, pro)*, *hsdΔ5*, *supE*, *thi*$^-$, F'(*lacI*^{q}, *lacZΔM15*, *pro*$^+$); Messing *et al.*, 1977) was used throughout for general purposes such as cloning, preparing templates for sequencing, etc.

2.1.4 Yeast Strains.

Table 2.1 Yeast Strains.

Strain	Genotype	Reference/Origin
S150-2B	**a**, *ura3-52, leu2-3,-112, trp1-289, his3-Δ1.*	Wellcome.
D150	**a**, *ura3-52, leu2-3,-112, ade1-100, his4-519.*	Tollervey *et al.*, 1991. Also known as BWG1-7A.
DBY745	α, *ura3-52, leu2, ade1-100.*	Legrain and Rosbach, 1989.
SPJ8.31	**a**, *prp8-1, ura3-52, leu2, his3.*	S. Jackson (this lab.).
JDY8.22	α, *prp8-2, ura3-52, leu2-3,-112, lys2, his3-Δ1.*	J. D. Brown (this lab.).
JDY8.31(I1)	α, *prp8-3, ura3-52, his3-Δ1.*	Ditto.
JDY8.31(I2)	α, *prp8-3, ade, his3-Δ1, leu2-3,-112, trp1-289, tyr, lys.*	Ditto.
JDY8.31(D1)	α, *prp8-3, his3-Δ1, leu2-3,-112, tyr, lys.*	Ditto.
JDY8.31(D2)	**a**, *prp8-3, ade, leu2-3,-112.*	Ditto.
JDY8.31(H2)	α, *prp8-3, ade.*	Ditto.
JDY8.57	α, *prp8-5, ura3-52, leu2-3,-112, ade1/2, his7, tyr1.*	Ditto.
DJY76	α, *prp8-7, ura3-52, leu2-3,-112, tyr1, his.*	D. Jamieson (this lab.).
JDY8.81	*prp8-8, ura3-52, leu2-3,-112, trp1-289, his3-Δ1.*	J. D. Brown (this lab.).
L149-7B	α, *dbf3-1, trp1-289, ura3-52.*	Shea *et al.*, 1994.
JL1U	**a**, *dna39-1, his7-2, trp1-289, ura3-52.*	Ditto.

Strain	Genotype	Reference/Origin
JL2U	**a**, *dna39-2, ade1, his7-2, trp1-289, ura3-52.*	Ditto.
JL3U	**a**, *dna39-3, ade1, his7-2, leu2-3, tyr1, trp1-289, ura3-52.*	Ditto.
JDY8.01	**a/α**, *ura3-52, leu2-3,-112, PRP8/prp8::LEU2, ade2-100/ADE2, HIS3/his3-Δ1, TRP1/trp1-289.*	Brown and Beggs, (1992).
JDY8.02	α, *prp8::LEU2, ura3-52, leu2-3,-112, ade2-100, his3-Δ1, trp1-289*; pY8000 (*URA3*).	Ditto.
JDY8.05	α, *PRP8::LEU2, ura3-52, leu2-3,-112, ade2-100, his3-Δ1, trp1-289*; pJDY13 (*HIS3*).	Ditto.
DJY105	α, *spp81-3, leu2-3,-112, ura3-52, his3-Δ1.*	D. Jamieson (this lab.).
J17	**a**, *prp9-1, ade1, ade2, ura1, his7, tyr1, lys2, gal1.*	J. Warner (New York).
J93	α, *prp9-1, ade, ura, leu, trp, arg.*	Ditto.
DJY85	**a/α**, *prp2-1, ura3, ade1/ADE1, ade2, trp1/TRP1, his3/HIS3, tyr1/TYR1, lys2-801/LYS2, can1/CAN1.*	D. Jamieson (this lab.).
MCY4	α, *LEU2-GAL1::uss1, ade1-101, his3-Δ1, trp1-289, ura3-52.*	Cooper *et al.*, 1995.
Y166	**a**, *cdc25-1, ade1, ade2, gal1, his7, lys2, tyr1.*	L. Johnston (NIMR, London).

2.1.5 Plasmid Vectors and Constructs.

Table 2.2 Plasmid Vectors and Constructs.

Plasmid	Description	Reference/Origin
pTZ18R, pTZ19R	Phagemid vectors with multiple cloning sites in *lacZ* α-peptide fragment. F1 origin allows synthesis of single-stranded DNA; *amp*^{r}.	Pharmacia.
pBluescript KS$^+$ and KS$^-$	Phagemid vectors with multiple cloning sites in *lacZ* α-peptide fragment. F1 origin allows synthesis of single-stranded DNA; *amp*^{r}.	Stratagene.
pRS316	Yeast plasmid based on pBluescript, low copy number; *ARS, CEN, URA3, amp*^{r}.	Sikorski and Hieter, 1989.
pBM125	Yeast plasmid, low copy number containing the *GAL1-10* promoter region (on a 0.81kb *Eco*RI-*Bam*HI fragment). Does not include *GAL1* ATG start codon; *ARS, CEN, URA3, amp*^{r}.	Johnson and Davis, 1984.
YEp24	Yeast plasmid, high copy number; *URA3*, 2μ, *amp*^{r}.	Botstein *et al.*, 1979.
pFL44S	Yeast plasmid based on pUC19, high copy number; *URA3*, 2μ, *amp*^{r}	Bonneaud *et al.*, 1991.
pY8000	YEp24 containing the entire *PRP8* gene on a 9.0kb insert.	Jackson *et al.*, 1988.
pY8500	YCp50 containing the entire *PRP8* gene on a 12kb insert; *CEN, ARS, URA3, amp*^{r}.	Ditto.
pJK3	*Sal*I fragment from pY8000, with *Nhe*I site engineered by mutagenesis, inserted into pBR322 with vector *Hin*dIII site destroyed.	Brown and Beggs, 1992.
pJDY6	pY8000 with engineered *Nhe*I site at N-terminus of *PRP8*.	Ditto.

Plasmid	Description	Reference/Origin
pJDY7	pJDY6 with region encoding the N-terminal acidic region of Prp8p deleted by *Nhe*I-*Xba*I fusion.	J.D. Brown (this lab).
pJDY23	Identical to pJDY6, except encodes *prp8-1* point mutant instead of wild-type *PRP8*.	Ditto.
pJDY10	pRS316 with entire *PRP8* gene and *GAL1-PRP8* chimeric promoter.	Brown and Beggs, 1992.
pJDY1	*Sal*I fragment from pY8000 inserted into pFL44.	J.D. Brown (this lab).
pJDY2	*PRP8 Sal*I to *Xba*I fragment from pY8000 inserted into pFL44.	Ditto.
pJDY3	*PRP8 Cla*I fragment from pY8000 inserted into pFL44.	Ditto.
pJDY4	*PRP8 Bgl*II to *Bam*HI fragment from pY8000 inserted into pFL44.	Ditto.
pRB327	YEp21 (*LEU2*, 2μ) containing entire *TUB1* gene (including intron) on a 3kb insert.	Schatz et al., 1986b.
pT7rp28s	Template for *in vitro* transcription with T7 polymerase; after linearization with EcoRI produces a transcript with most of E1, IVS and E2.	M. Lossky, (this lab.).
cDNA expression library	Library was constructed by directional cloning of yeast cDNAs downstream of the *GAL1* promoter region in a pRS316-based vector.	Liu *et al.*, 1992.
pBM-PRP2	Encodes wild-type *PRP2* under transcriptional control of the *GAL1* promoter. Derived from pBM125.	King and Beggs, 1990.
pBM-PRP2^{dn1}	Identical to pBM-PRP2 except encodes dominant negative point mutant *PRP2LAT*.	Plumpton *et al.*, 1994.
pSM30	pTZ18R containing yeast *DED1* gene. A 2.5kb *Bam*HI-*Sal*I fragment was used to probe Northern blots.	S. Maccallum, (this lab.).
pDP6	pUC9 containing the yeast *LYS2* gene. A 5kb *Hin*dIII	Fleig *et al.*, 1986.

Plasmid	Description	Reference/Origin
	fragment was used to probe Northern blots.	
pBRCYH2	pBR322 containing yeast *CYH2* gene. An *Eco*RI fragment was used to probe Northern blots.	M. Lossky, (this lab.).
pYA301	pBR322 containing the yeast *ACT1* gene. A 1.0kb fragment (*Bam*HI-*Bgl*II) was used to probe Northern blots.	Gallwitz and Sures, 1980.

2.1.6 Synthetic Oligodeoxynucleotides.

All oligodeoxynucleotides were synthesised at the OSWEL DNA service, Edinburgh, except CR1, which was synthesised by Caroline Russel (this laboratory) on a Beckman 1000 DNA synthesizer.

Table 2.3 Oligodeoxynucleotides.

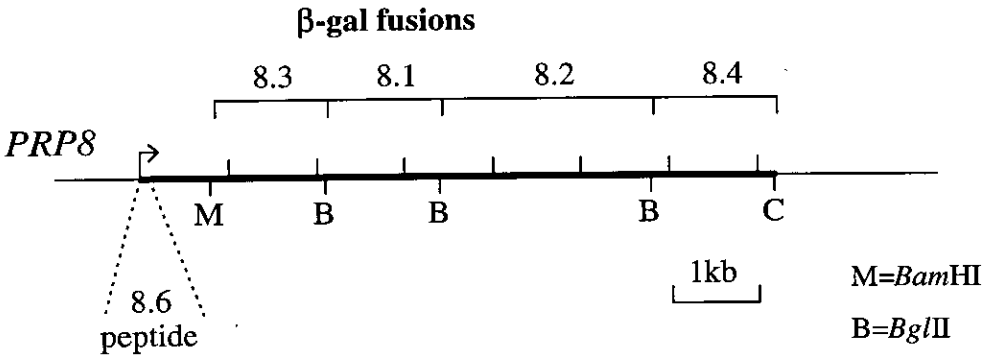
Name	Sequence (5' to 3')	Description
486A	CTT AAG GTA AGT AT	Complementary to 5' end of yeast U1 snRNA.
626C	CTA CAC TTG ATC TAA GCC AAA AGG C	Complementary to an internal region of yeast U2 snRNA.
483A	CCG TGC ATA AGG AT	Complementary to the 5' end of yeast U4 snRNA.
485A	AAT ATG GGC AAG CCC	Complementary to an internal region.of yeast U5 snRNA.
6A	TCA TCT CTG TAT TG	Complementary to an internal region of yeast U6 snRNA.
A969	GGG GGG AAG CTT CTC TAG ACC wGG xGG yGG zGG w=A or T; x=A or C y=A or G or T; z= G or T.	Complementary to any copy of the repetitive LP_nG-encoding motif at N-terminus of *PRP8*; contains an *Xba*I site. To create partial deletion mutants. Degenerate for 2 or 3 nucleotides at four positions as shown (lower case).
989S	GGG GGG GGG AAG CTT CTC	Complementary to A969; contains *Xba*I site.
819X	TCA AAG GGT GTG GGC CG	Complementary to PRP8 positions 4528 to 4544, orientation downstream.
M5583	AGA TAT CCT GGT GAA AGG	Complementary to PRP8, used to amplify the region containing the *dbf3-1* mutation

Name	Sequence (5' to 3')	Description
M5584	GTC CAC CAA AAG GGG CTG	from L149-7B genomic DNA.
G0260	TTG GGC AGT ATA TAC AGG	Complementary to *PRP8* positions 648 to 631; orientation upstream.
A153	GTC GAC TTT AGA GCG CCT AG	Partial complement of *PRP8* (positions 499 to 484); orientation upstream. Contians *Sal*I site.
229G	CAC TCG AGC TGA GCT GTC ATT TAT ATT G	Partial complement of a region of *GAL1-10* promoter 100bp upstream of *GAL10* protein start; oriented towards *GAL1* transcriptional start. Contains *Xho*I site.
024R	CCA CTT TAA CTA ATA CTT TC	Complementary to *GAL1-10*. Anneals just upstream of *GAL1* transcriptional start; orientation towards *GAL1*.
G0879	TTT CGG TTT GTA TTA CTT C	Complementary to *GAL1-10*. Anneals just upstream of *GAL1* transcriptional start; orientation towards *GAL1*.
P0537	TTG GGC TCT AAA TCC ACG TA	Complementary to either *TUB1* or *TUB3* yeast genes, to region downstream of intron. Orientation upstream.
P3333	ACC CAC GCG TAC CCA AGA TCT G	Partial complement of a region just upstrem of the *TUB1* start codon. Used to amplify by PCR (with P0537) a region containing the *TUB1* intron. Contains *Mlu*I site.
762N	TAA TAC GAC TCA CTA TA	T7 Promoter primer.
976N 977N	CTA GCC CAG ACA TTG TGT TGC GCG CG GAT CCG CGC GCA ACA CAA	Anneal to form linker with *Bam*HI and *Nhe*I ends, containing *Bss*HII site, ATG start codon and yeast translational start sequence. Used to fuse *GAL1-10* to *PRP8*.

Name	Sequence (5' to 3')	Description
	TGT CTG GG	
575S 453S	CGT CTA TTT CAT CTA GAT GAA ATA GA	Anneal to form linker with *Xba*I and *Cla*I ends. To make in-frame deletion of most of *PRP8* gene.
M0291 M0292	CTA GAC CGG CCG CTA GCC CAG ACA TTG TGT TG GAT CCA ACA CAA TGT CTG GGC TAG CGG CCG GT	Anneal to form linker with *Xba*I and *Bam*HI ends, containing *Nhe*I and *Eag*I sites, ATG start codon and yeast optimal translational start sequnce. To introduce unique *Eag*I site in *prpΔN*.
N2437	GGC CAG AAG CTT CT	Self-complementary. To destroy *Eag*I site and introduce *Hin*dIII.
N6930 N6931	GGC CGC GGC TCC CGG AGG TGG CGG CGG TGG AGG GGG TGG CGG TAA GGC CTT ACC GCC ACC CCC TCC ACC GCC GCC ACC TCC GGG AGC CGC	Anneal to form linker with *Eag*I ends, containing *Sac*II site. To create in-frame insertions into *prp8ΔN*, encoding polyproline in one orientation and polyglycine in the other.
N7684 N7685	GGC CGC GAG CTC TTC GTC CTC ATC CTC GTC CTC TTC ATC GGC CGA TGA AGA GGA CGA GGA TGA GGA CGA AGA GCT CGC	Anneal to form linker with *Eag*I ends, containing *Sac*I site. To create in-frame insertions into *prp8ΔN*, encoding acidic amino acids in one orientation and polyserine in the other.
CR1	GGC CAA CAC CCG GGT GTT	Self-complementary, anneals to form linker with *Eag*I ends. To create in-frame insertions into *prp8ΔN* encoding amino acid sequence ANTRVL.

2.1.7 Antisera and Antibodies.

Antibodies have been raised against β-galactosidase-Prp8 fusion proteins 8.1 to 8.4 (Lossky *et al.*, 1987; Jackson *et al.*, 1988), and also against a synthetic 35 amino acid pepetide (anti-8.6 antibodies; Brown, 1992). Regions of *PRP8* incorporated in the β-galactosidase fusions are indicated in Figure 2.1. Amino acids 2 to 35 of the 8.6 peptide are identical to amino acids 2 to 35 of Prp8p.



Sequence of 8.6 peptide:

CSGLPPPPPGFEEDSDLALPPPPPPPPGYEIEE

Figure 2.1: Regions of *PRP8* against which antibodies have been raised. Thick line: *PRP8* coding sequence.

2.1.8 Size Markers.

Size markers for nucleic acid and protein electrophoresis were as follows: 1kb ladder markers (Gibco BRL, catalogue number 15615-016), or bacteriophage λ DNA (Gibco BRL, catalogue number 25250-010) digested with *Bst*EII or *Hin*dIII were used for agarose gel electrophroesis of DNA. Sigma pre-stained protein markers (catalogue number SDS-7B) or high molecular weight markers (SDS-6H) were used for SDS-polyacrylamide gel electrophoresis of proteins.

2.1.9 Buffers.

10xTAE:	0.4M Tris-acetate (pH7.5), 20mM EDTA.
10xTBE:	0.9M Tris-borate (pH8.3), 20mM EDTA.
10xTB:	0.9M Tris-borate (pH8.3).
10xTBS:	0.5M Tris-HCl, 1.5M NaCl (pH7.5).
20x SSC:	3M NaCl, 0.3M tri-sodium citrate (pH7.0).

2.2 GENERAL METHODS.

2.2.1 General Guidelines.

The following guidelines were observed throughout, except where indicated otherwise. Procedures were carried out at room temperature. Small-scale procedures were carried out in 1.5ml sterile microfuge tubes (Treff), except polymerase chain reactions which were carried out in sterile, thin-walled 0.5ml tubes (Treff). Larger scale procedures were carried out in 30ml polypropylene tubes, sterile 15ml or 30ml Corex tubes, 10ml Sterilin tubes or 50ml Greiner tubes. For harvesting large volume cultures, Sorvall 250ml or 500ml pots were used.

Small volumes (up to 1.5ml) in Treff tubes were centrifuged in an Eppendorf benchtop microcentrifuge at 14,000 rpm. Centrifugation of volumes greater than 50ml was carried out in a Sorvall Superspeed centrifuge using either a GS-A or GS-3 rotor at 5000 rpm. Intermediate volumes (1.5ml to 50ml) were centrifuged in either SS34 or HB-4 rotors in the Sorvall Superspeed centrifuge, or in an IEC CENTRA-4X bench-top centrifuge at 4400 rpm in a 215A swing-out rotor. Ultracentrifugation was performed in a Sorvall OTD50, 55 or 65B ultracentrifuge (Du Pont).

2.2.2 Sterilization Procedures.

Plastics (such as microfuge tubes) and solutions were sterilized by autoclaving at high temperature and pressure (120°C and 15 pounds per square inch) for 20min. Solutions prepared in small volumes or from heat-labile materials (DTT, tryptophan, SDS, and antibiotics) were sterilized by filtration through disposable 0.45 micron Acrodisc filter units

(Gelman Sciences). To inactivate ribonucleases Corex tubes, glass pipettes, and other glassware to be used in work with RNA was sterilized by baking for 250°C for 16hr.

2.2.3 Deionization of Solutions.

Formamide and solutions of acrylamide were deionized by mixing with 0.1 volume of analytical mixed-bed resin (20-50 mesh; Bio-Rad Laboratories), stirring for 30mins. The resin was removed by filtration through Whatman No.1 filter paper. Acrylamide and formamide were stored at 4°C.

2.2.4 Photography.

After electrophoresis on agarose gels containing 1μg/ml ethidium bromide, nucleic acids were visualised by transillumination under short wavelength ultraviolet light (254nm) and photograghed using the GDS 5000/UP-860CE digital camera and gel documentation system (Ultra Violet Products and Sony).

2.2.5 Autoradiography.

[^{32}P]-labelled nucleic acids in gels or on filters were detected by exposure to Cronex X-ray film (Du Pont) in a light-proof cassette with a calcium-tungstate intensifying screen at -70°C. [^{35}S]-labelled DNA was also detected by exposure to X-ray film, but at room temperature and without an intensifying screen.

2.2.6 Phosphorimaging.

Storage Phosphor screens (Molecular Dynamics) were erased by exposure to light for 6mins, and exposed to thoroughly wrapped filters in the cassette provided by the manufacturer. After 1-48hr exposure, the screen was scanned in the Phosphorimager (Molecular Dynamics) and signals quantified using the Imagequant software package. Data was stored on 80Mbyte RHOMAT magnetic tapes (3M).

2.3 MICROBIOLOGICAL METHODS.

2.3.1 Propagation and Storage of *E.coli.*

Growth and storage of *E.coli* were carried out as described in Sambrook *et al.*, (1989).

2.3.2 Transformation of *E.coli.*

Competent cells were prepared by a modification of the method of Hanahan (1983).

1) 100-500ml of LB was inoculated with 1/100th volume of a fresh stationary culture of strain NM522 and grown at 37°C to an OD_{600} of 0.3-0.4.

2) The culture was chilled on ice for 15-30 mins. Cells were harvested at 4°C, and resuspended in ice-cold 75mM $CaCl_2$ (50% original culture volume), and incubated at 0°C for 30mins.

3) Cells were harvested, again at 4°C, resuspended in TFBII (4% original culture volume), and incubated on ice for 30mins.

4) Cells suspended in TFBII were transformed immediately, or were snap-frozen in 0.5ml aliquots in liquid nitrogen and stored for 1-2 months at -70°C. Frozen competent cells were thawed on ice prior to transformation.

5) For transformation, 50-200µl aliquots of competent cells were mixed with DNA and incubated on ice for 15min. High concentrations of DNA inhibit transformation, so less than 10ng of plasmid or a fraction of a ligation reaction was used.

6) Cells were heat-shocked to 42°C in a water-bath for 90secs, and placed on ice for 5mins.

7) 1ml LB was added and incubated at 37°C for 1hr to allow expression of the ampicillin resistance gene to commence.

8) 100µl was spread to a 10cm LB-amp plate and grown overnight at 37°C.

9) Plates containing IPTG and X-Gal (50 and 60µg per 10 cm plate, respectively) were incubated at 4°C to allow colonies expressing β-galactosidase (non-recombinants) to develop a blue rather than white (recombinant) colour.

Buffers:

TFBII:- 10mM MOPS (pH6.8, adjusted with KOH), 10mM RbCl, 75 mM $CaCl_2$, 15% (v/v) glycerol. Sterilise by autoclaving.

2.3.3 Propagation and Storage of Yeast.

Basic methods for the storage and propagation of yeast are described in Guthrie and Fink (1991).

2.3.4 Sporulation and Spore Recovery.

Diploid strains to be sporulated were streaked to rich medium (YPDA) and incubated at 30°C for at least 16hr. Proliferating cells were recovered by scraping and resuspended in 3ml liquid sporulation medium in a sterile test tube to a density of about 0.5 OD_{600}. Sporulating cells were incubated at 23°C on a turning wheel until asci were visible under light microscopy (3 to 7 days, variable depending on the strain). Thereafter, aliquots of this culture were harvested by gentle centrifugation (5000 rpm in benchtop microfuge), resuspended in 90µl buffer (1.2M sorbitol, 10mM EDTA, 100mM citric acid) to give a slightly turbid suspension, and 10µl β-glucuronidase (Boeringer Mannheim) added to partially digest the ascus wall. Dissection of asci was performed on a Singer MSM microdissector (Singer Instrument Co.). Discrete, dissected spores were placed on YPDA agar and incubated at 23°C until colonies formed.

2.3.5 Transformation of Yeast.

Yeast transformation was carried out according to the basic protocol of Ito *et al.* (1983) with the modifications of Gietz *et al.* (1992).

1) A culture of yeast was grown overnight in an appropriate medium to OD_{600} 0.8.

2) An aliquot of this culture was diluted 10- to 12-fold into 100ml of fresh, prewarmed medium (to 0.06 to 0.07 OD_{600}), and grown to 0.3 OD_{600}.

3) Cells were harvested, washed with sterile distilled water and resuspended in 1ml water, and transferred to microfuge tubes.

4) Cells were harvested and resuspended in TE/LiAc to 60-70 OD_{600}/ml.

5) 50µl cell suspension was mixed with 1µg plasmid DNA, 5µg denatured sonicated salmon sperm DNA (10mg/ml) and 300µl PEG buffer, and the mixture incubated at 30°C for 30min on a turning wheel.

6) The mixture was heat-shocked to 42°C for 15mins.

7) Tubes were spun in a microfuge for 3min, and the cell pellet resuspended in 1ml sterile water. 100µl aliquots were spread to plasmid-selective plates.

Buffers:

TE/LiAc:-	1ml 10xTE, 1ml 10xLiAc, 9ml sterile distilled water.
PEG buffer:-	1ml 10xTE, 1ml 10xLiAc, 9ml 50% (w/v) PEG-4000.
10xTE:-	0.1M Tris-HCl pH7.5, 0.01M EDTA.
10xLiAc:-	1M LiAc, adjusted to pH7.5 with acetic acid.

2.4 NUCLEIC ACID METHODS.

The following standard manipulations of RNA and DNA were carried out according to Sambrook *et al.* (1989) or according to equipment manufacturers instructions: quantitation by spectrophotometry or ethidium bromide fluorescence, deproteinization by phenol and phenol/chloroform extraction, alcohol precipitation, restriction endonuclease digestion and agarose gel electrophoresis of DNA, dephosphorylation of DNA and ligation of DNA fragments. Solutions of circular, plasmid DNA in TE (10mM Tris-HCl pH 8.0, 1mM EDTA) were stored at 4°C. Linear DNA fragments and single-stranded DNA were stored at -20°C. Solutions of RNA in sterile, double-distilled water were stored at -70°C. Solutions used in the analysis of RNA were treated with diethyl pyrocarbonate (DEPC) to inactivate RNase. 0.1% (v/v) DEPC was added to the solution and stirred for 1hr, prior to autoclaving. Solutions containing chemicals labile to DEPC treatment (such as Tris) were not treated directly, but made up with DEPC-treated dH_2O.

2.4.1 Plasmid DNA Preparations.

2.4.1.1 Minipreps.

Small scale preparations of DNA were based on the method of Birnboim and Doly, (1979).

1) 5ml LB-amp was inoculated from a single *E.coli* colony and incubated overnight at 37°C with shaking.

2) 1.5ml of this culture was harvested in a microfuge for 5mins.

3) The culture medium was discarded and the cells resuspended in 100µl solution A.

4) 200µl solution B was added, and the cells lysed by gentle inversion and incubation on ice for 5mins.

5) 150µl solution C was added and mixed by inversion, and placed on ice for a further 5mins.

6) Debris and chromosomal DNA was removed by centrifugation for 5mins in a benchtop microfuge and the supernate transferred to a fresh tube.

7) The supernate was extracted once with phenol/chloroform and nucleic acids precipitated with ethanol and sedimented by centrifugation in a benchtop microfuge at 4°C for 10mins.

8) The pellet was rinsed in 70% (v/v) ethanol, dried *in vacuo*, and dissolved in 50µl TE. To remove RNA, RNase A (Boeringer Mannheim) was added to restriction digests to a final concentration of 0.5mg/ml.

Buffers:

A:- 50mM glucose, 25mM Tris-HCl (pH8.0), 10mM EDTA.

B:- 0.2M NaOH, 1% (w/v) SDS, prepared immediately before use.

C:- 3M sodium acetate pH5.3.

2.4.1.2 Midipreps.

Medium scale preparations of plasmid DNA followed a similar protocol to the miniprep method above.

1) 50ml LB-amp was inoculated from a stationary-phase culture and grown overnight.

2) Cells were harvested in a 50ml Greiner tube and resuspended in 2ml of solution A.

3) 100µl freshly-prepared 20mg/ml lysozyme in solution A was added, and incubated for 5mins on ice.

4) 4ml solution B was added, and mixed by inversion until the solution became clear.

5) 3ml solution C was added, and the lysate mixed and incubated on ice for 30mins.

6) Debris was sedimented in an SS34 rotor (4°C, 10k rpm, 10mins). The supernate was transferred to a 30ml Corex tube, and 18ml ice-cold ethanol added.

7) Precipitated DNA was sedimented by centrifugation in an HB-4 swing-out rotor (4°C, 10k rpm, 10mins). The pellet was dissolved in 2ml solution D.

8) After one phenol/chloroform extraction, and one back-extraction again to 2ml solution D, each aqueous phase was pooled in a 15ml Corex tube.

9) 8ml of ice-cold ethanol was added incubated at -20°C for 20mins to precipitate nucleic acids.

10) After sedimentation, the pellet was dissolved in 0.5ml TE.

11) RNase A was added to a concentration of 0.2µg/ml and incubated for 1hr at 37°C.

12) Extraction was performed once with phenol, twice with phenol/chloroform and lastly once with chloroform.

13) Plasmid DNA was precipitated with two volumes ice-cold ethanol and 0.1 volume 4M NaCl, and the pellet washed with 70% (v/v) ethanol, dried *in vacuo* and dissolved in 0.2-0.4ml TE.

This procedure routinely yielded 400-500µg supercoiled plasmid DNA.

Buffers:

A, B, and C:- As 2.4.1.1 above.

D:- 0.1M sodium acetate pH5.2, 1mM EDTA, 0.1% (w/v) SDS, 40mM Tris-HCl (pH8.0).

2.4.1.3 Preparation of cDNA Library.

The midiprep method above was modified to prepare library DNA, to ensure maximum recovery of plasmid DNA without distortion of the library which can accompany amplification. The library was supplied in the form of *E.coli* transformants on agar, which were recovered by scraping, resuspended in sterile dH_2O, and spread onto LB-amp plates. After 16hr incubation at 37°C, when colonies had grown to an almost-confluent layer, the transformants were recovered from the surface of the agar by scraping. Half of these amplified cells were resuspended in LB containing 40% glycerol and snap-frozen for storage at -70°C. The remainder were resuspended in two 50ml volumes LB-ampicillin, and after 6 hours incubation at 37°C with shaking plasmid DNA was prepared according to the midiprep method (section 2.4.1.2).

2.4.1.4 Maxipreps.

To purify large quantities of plasmid DNA, the method of centrifugation to equilibrium on caesium chloride density gradients was used. This procedure separates DNA from RNA without the use of RNase A, and was used where the purity of DNA, especially the absence of contaminating RNase, was important (for example to prepare templates for *in vitro* transcription). The alkaline lysis method was based on that of Maniatis *et al.*, (1982).

1) 500ml LB-amp was inoculated from a fresh stationary culture and grown overnight at 37°C to stationary phase.

2) Cells were chilled on ice and harvested at 4°C in 250ml Sorvall pots in a GS3 rotor (5k rpm, 10mins).

3) The cell pellet was resuspended in 7.5ml 25% sucrose, 50mM Tris-HCl (pH7.5), transferred to a 50ml Greiner tube and placed on ice for 10mins.

4) 2.5ml lysozyme (10mg/ml in dH_2O, freshly prepared) was added and incubated for 30mins.

5) 12ml lysis buffer (0.1% (v/v) Triton, 60mM EDTA, 50mM Tris-HCl, pH8.0) was added, mixed gently and incubated for 10mins on ice.

6) Debris was removed by centrifugation (12k rpm, 45mins), and the supernate transferred to a fresh 50ml Greiner tube.

7) Finely ground caesium chloride (1g per 1ml supernate) was added and dissolved by gentle agitation.

8) 3ml 5mg/ml ethidium bromide was added, and the solution centrifuged overnight in a TV865B rotor (45k rpm, 20°C) in plastic ultracentrifuge tubes.

9) DNA was visualised by exposure to shortwave UV light from a transilluminator and the supercoiled plasmid DNA band (lower band) was removed (in approximately 3ml) with a large-bore hypodermic needle.

10) Ethidium bromide was removed by extraction with water-saturated, CsCl-saturated propan-2-ol until no trace of the red colour remained in the aqueous phase (4-6 extractions).

11) 8ml TE and 8ml propan-2-ol was added to the aqueous phase in a fresh 30ml Corex tube, and incubated at -20°C for 30mins. This step precipitates DNA but leaves most of the CsCl in solution.

12) Plasmid DNA was sedimented by centrifugation in an HB-4 swingout rotor (10k rpm, 4°C, 30mins) and the pellet washed and redissolved in 0.5ml TE. DNA produced by this method is stable for long periods at -20°C, but may gradually degrade at 4°C due to contaminating DNase from the bacterial lysate. These may be removed by phenol/chloroform extraction followed by precipitation.

This procedure routinely yielded more than 500µg supercoiled plasmid DNA.

2.4.2 Sequencing.

All sequencing reactions were carried out using the Sequenase version 2.0 kit (USB) according to the manufacturers instructions. Templates were either preparations of single-stranded phagemid DNA, or single-stranded or double-stranded DNA produced by the polymerase chain reaction (PCR). Products of sequencing reactions were fractionated by denaturing electrophoresis on 6% (w/v) polyacrylamide/8M urea gels in 1xTB buffer, fixed in 10% (v/v) methanol/10% (v/v) acetic acid and dried (Bio-Rad vaccuum drier).

2.4.2.1 Preparation of Phagemid DNA (Single-stranded).

1) Cultures were grown in large (250ml) flasks with vigorous shaking to ensure good aeration. 10ml 2xYT-amp was inoculated with 200µl of a fresh stationary culture of

E.coli bearing a sequencing plasmid (ie pTZ- or pBluescript-derived), and incubated at 37°C.

2) After 30mins, 10µl M13KO7 helper phage (Pharmacia) was added, and after a further 1hr 14µl 50mg/ml kanamycin (to select for infection by helper phage) added, and incubation continued for 5-6hr.

3) Cells were pelleted in microfuge tubes and the supernate transferred to fresh tubes.

4) 10% volume PEG/NaCl was added and incubated at room temperature for 20mins.

5) Phage were sedimented by centrifugation and the supernate discarded. Tubes were spun for a further 2min and the residual supernate also discarded, as contaminating PEG interferes with the sequencing reaction.

6) The pellet was resuspended in 100µl dH_20 and extracted twice with phenol/chloroform.

7) DNA was precipitated with ethanol and the pellet washed in 90% (v/v) ethanol and dried *in vacuo*.

8) The pellet was dissolved in 20µl sterile dH_20. Typically 2µl was electrophoresed to confirm yield of single-stranded DNA (typically 1-2µg) and 7µl used in the Sequenase reaction.

PEG/NaCl: 20% (w/v) PEG-6000, 4M NaCl.

2.4.2.2 Production of Single-stranded DNA by Primer Extension with Thermostable Polymerase.

This method is modified from supplement 16 of Ausubel *et al.*, (1987).

1) The standard PCR reaction (section 2.4.13) was carried out and product purified on a Quiquick Spin Column (Quiagen).

2) A second PCR reaction (100µl) was assembled with only one primer (40pmol of one of the primers used in step 1). 5µl purified PCR product (10% of total product) from step 1 was used as template.

3) The reaction volume was overlaid with mineral oil and subjected to 25 cycles (92°C 30sec, 50°C 30sec, 72°C 1-5mins) in a Hybaid Thermal Reactor. In each cycle the primer is extended and single-stranded DNA accumulates.

4) After the final extension step, the mineral oil was removed by pipetting, and 50µl sterile dH_2O added.

5) 150µl 5M ammonium acetate and 300µl of ethanol (room temperature) were added, and incubated for 5mins at room temperature.

6) DNA was sedimented and the pellet washed with 70% (v/v) ethanol (room temperature) and dissolved in 20µl sterile dH_2O. 7µl was used in the Sequenase reaction.

2.4.2.3 Preparation of DNA for Direct, Double-stranded Sequencing.

This method is modified from supplement of 16 Ausubel *et al.*, (1987).

1) A PCR reaction was carried out according to the standard protocol (section 2.4.13) and the product purified on a Quiquick Spin Column (Quiagen).

2) To 1µl of purified PCR product was added 8µl dH_2O and 2µl sequenase reaction buffer.

3) Denaturation was carried out in a heated aluminium block (95°C) for 3mins, after which tubes were directly transferred to liquid nitrogen and snap-frozen for 1min.

4) 1µl sequencing primer was added to the wall of the microfuge tube above the frozen pellet, and the tube spun for 2min until the pellet had thawed.

5) The mixture was incubated at room temperature for 30mins to allow the primer to anneal, and transferred to ice until sequencing. This volume was used in the extension step of the Sequenase reaction.

2.4.3 Extraction of Genomic DNA from Yeast.

This procedure yielded DNA of sufficient quality for restriction digestion and Southern blotting, or use as a template in the polymerase chain reaction. Manipulations of genomic DNA were carried out with disposable pipette tips cut down to give a wider aperture, to minimise shearing forces.

1) 20ml of an appropriate medium was inoculated and grown overnight.

2) Cells were harvested and resuspended in 400µl sorbitol/$NaPO_4$.

3) 100µl lyticase at 2500U/ml in sorbitol/$NaPO_4$ was added, and cells incubated at 37°C for 1hr on a turning wheel.

4) Sphaeroplasts were pelleted by gentle centrifugation (in a microfuge at half-speed) and resuspended in 0.5ml TE.

5) 50µl 20% (w/v) SDS was added, mixed by inversion and incubated at 65°C for 30min.

6) 200µl 5M potassium acetate (pH5.7) was added and incubated on ice for 1hr.

7) Cell debris was removed by spinning in a microfuge for 5min, and the supernate transferred to a fresh tube.

8) Genomic DNA was precipitated by addition of an equal volume of propan-2-ol, and incubation at room temperature for 5min.

9) Precipitated DNA was recovered by spooling onto a sealed Gilson pipette tip and redissolved in 300µl TE.

Buffers:

Sorbitol/$NaPO_4$:- 1.2M sorbitol, 50mM $NaPO_4$ (pH7.0), 5mM DTT, prepared immediately before use.

TE:- 50mM Tris-HCl (pH8.0), 20mM EDTA.

2.4.3.1 Plasmid Recovery from Yeast to *E.coli.*

The above method was adapted to extract DNA from yeast to recover plasmids for transformation of *E.coli*. DNA prepared by this method gives a greater efficiency of transformation than DNA prepared as above (Plumpton, 1993).

1) to 4) as for 2.4.3 above.

5) The suspension of sphaeroplasts was vortexed vigorously, and debris removed by centrifugation.

6) To the supernate, 50µl 20% (w/v) SDS was added, and incubated at 65°C for 30min.

7) 200µl solution B (0.2M NaOH, 1% (w/v) SDS) was added and mixed.

8) 150µl 5M potassium acetate was added and incubated on ice for 10min.

9) Debris was removed by centrifugation, and the supernate mixed with an equal volume of propan-2-ol in a fresh tube.

10) Precipitated material was sedimented and redissolved (without drying) in 200µl dH_2O, and phenol/chloroform extracted.

11) DNA was precipitated with ethanol and redissolved in 50-100µl dH_2O. Aliquots were used to transform *E.coli* of high competence.

2.4.4 Extraction of RNA from Yeast.

1) To induce temperature-sensitive splicing defects, logarithmically growing cultures were diluted with identical volumes of the same medium preaerated and prewarmed to the appropriate temperature. For example, to heat-shock a 100ml YPDA culture from 23°C to 37°C, 100ml YPDA prewarmed to 51°C was added. For controls cultures were diluted without changing the temperature. Incubation was continued at the temperature resulting from heat shock.

2) 200ml volumes of culture were transferred to 250ml Sorvall pots and chilled on ice for 15min.

3) Cells were harvested in a GSA rotor (5k rpm, 10min, 4°C), resuspended in 10ml ice-cold TNE, and transferred to a 50ml Greiner tube.

4) Cells were sedimented and the supernate discarded. The cell pellet was snap-frozen by immersion of the tube in liquid nitrogen. Frozen cell pellets were either treated to extract RNA immediately or first stored for up to 48hr at -70°C.

5) 1ml TNE, 0.3ml glass beads (BDH, 100 mesh) and 1ml phenol were added to the frozen cell pellet, and the mixture vortexed vigorously until the pellet thawed and a further 90sec thereafter.

6) 3ml TNE, 0.2ml 20% (w/v) SDS, and 3ml phenol were added and vortexed again for 1min.

7) The aqueous phase was separated by centrifugation and transferred to a fresh 50ml Greiner tube.

8) Two phenol/chloroform extractions were carried out, and the aqueous phase transferred to a 15ml Corex tube.

9) RNA was precipitated by addition of 2.5 volumes ice-cold ethanol, and incubation on ice for 30min.

10) RNA was sedimented by centrifugation in an HB-4 swing-out rotor (4°C, 12k rpm, 20min), and the pellet dried in air and dissolved in 200µl DEPC-treated dH_2O.

This procedure routinely yielded 1-3mg RNA.

Buffers, etc:

TNE:- 50mM Tris-HCl (pH7.5), 100mM NaCl, 5mM EDTA.

Glass beads were prepared by boiling in 1M HCl for 30min and washing extensively with dH_2O, followed by baking at 250°C for 16hr.

2.4.5 Labelling Restriction Fragments by Random Priming.

This procedure is modified from Feinberg and Vogelstein (1984).

1) A restriction digest was fractionated on an agarose gel and a DNA fragment recovered using the Quiex kit (Qiagen) according to the manufacturers instructions. The DNA fragment was dissolved in dH_2O at 10-100ng/µl, and stored at -20°C.

2) 50-100ng of fragment was made up to 35µl with dH_2O and denatured by boiling for 3min and then cooled on ice.

3) 10µl OLB, 1µl BSA (20mg/ml), 3µl [α-^{32}P]dCTP (10mCi/ml), and 2 units *E.coli* DNA polymerase Klenow fragment were added and incubated at 37°C for 1hr.

4) Unincorporated nucleotides were removed by passage over a NAP-5 gel filtration column (Pharmacia) equilibrated with 10mM sodium phosphate (pH7.0).

Buffers:

OLB was prepared from the following solutions:

O: 1.25M Tris-HCl (pH8.0), 0.125M $MgCl_2$.

A: 1ml solution O, 18µl β-mercaptoethanol, 5µl each of dATP, dGTP, dTTP, each at 0.1M.

B: 2M HEPES (pH6.6, adjusted with 5M NaOH)

C: Random hexamer deoxyribonucleotides (Pharmacia) at 90 OD units/ml in TE.

OLB is a mixture of A:B:C in the ratio 2:5:3. OLB was stored at -20°C.

2.4.6 Labelling Oligodeoxynucleotides at the 5'-end.

Northern blots of snRNAs were probed with oligodeoxynucleotides, which were radiolabelled by addition of a ^{32}P-phosphate group to the 5'-hydroxyl using T4 polynucleotide kinase.

1) The reaction was assembled: 20pmol oligodoexynucletide, 3µl 10x PNK buffer (supplied by manufacturer), 20µCi [γ-^{32}P]ATP, sterile dH_2O to 30µl.

2) 5U T4 polynucleotide kinase (10U/µl, New England Biolabs) was added and incubated at 37°C for 1hr.

3) To probe for several snRNAs in a single hybridisation, separate labelling reactions were pooled and made up to 0.5ml with 10mM sodium phosphate (pH7.0). Unincorporated nucleotides were removed by passage over a NAP-5 gel filtration column (Pharmacia) equilibrated with 10mM sodium phosphate (pH7.0).

2.4.7 Denaturing Agarose Gel Electrophoresis of RNA.

RNA was size-fractionated by denaturing electrophoresis through formaldehyde/agarose gels, using a MOPS-containing buffer. Gels contained ethidium bromide to allow visualisation of ribosomal RNA.

1) For a 1.4% gel, 4.2g agarose (or 3g agarose for a 1% gel) was added to 161ml DEPC-treated dH_2O and 30ml 10x MOPS buffer, and boiled for 3-5min until the agarose had dissolved.

2) After the agarose solution had cooled to about 65°C, 15.3ml 37% (w/v) formaldehyde and 15µl ethidium bromide (10mg/ml) were added and mixed by swirling, and the 20cm x 25cm gel poured *in situ*.

3) For double wells (having combined two teeth of the comb with autoclave tape), 60µg yeast total RNA was made up to 50µl with DEPC-treated dH_2O and 50µl FSB added. Alternatively for single wells, 25µg RNA was made up to 20µl and 20µl FSB added. Samples were denatured by heating in a waterbath at 65°C for 5min, and then cooled on ice.

4) The apparatus was filled with 1x MOPS buffer up to just below the surface of the gel (ie the gel was not submerged), and samples loaded and run approximately 1cm into the gel. On an outer lane, gel loading buffer was run as a marker.

5) The gel was submerged under the buffer and run at up to 10V/cm. Buffer was recirculated with a peristaltic pump.

6) After electrophoresis the gel was wrapped in Saran wrap and photographed on a transilluminator.

The quantities in steps 1) and 2) above were scaled down by half to pour an 11cm x 14cm gel.

Buffers:
FSB:- 100µl 10x MOPS, 200µl deionized formamide, 120µl formaldehyde (37%, w/v).
Gel loading buffer: 0.25% bromophenol blue, 0.25% xylene cyanol FF, 15% Ficoll.
10x MOPS: 0.2M MOPS, 0.05M sodium acetate, 0.01M EDTA adjusted to pH7.0 with NaOH.

2.4.8 Denaturing Polyacylamide Gel Electrophoresis of Nucleic Acids.

DNA and RNA were size-fractionated under denaturing conditions on polyacrylamide/8M urea gels. For DNA, gels contained TB buffer (1x) and for RNA TBE buffer (1x). For general purposes, 6% polyacrylamide/8M urea (sequencing) gels were used; with shorter length fragments (less than 50 nucleotides), 10% acrylamide gels were used.

2.4.9 Capillary Blotting of RNA and DNA.

Southern and Northern blotting of agarose gels containing DNA and RNA respectively was performed by drawing 10x SSC through the gel into a stack of paper towels by capillary action, as described in Sambrook *et al.* (1989). DNA and RNA were blotted to Hybond-N nylon membrane (Amersham). Prior to transfer, gels were treated as follows:
(i) RNA gels were washed twice in 10x SSC, each for 20min.
(ii) DNA gels were washed twice in denaturation buffer (0.5M NaOH, 1.5M NaCl; each for 15min) and then twice in neutralisation buffer (0.5M Tris pH7.4, 1.5M NaCl; each for 30min).

After transfer, nucleic acids were crosslinked to the membrane by exposure to UV light in a UV Stratalinker 1800 (Stratagene).

2.4.10 Electroblotting of RNA from Polyacrylamide Gels.

After fractionation on 6% polyacrylamide/8M urea gels, RNA was transferred electrophoretically to Hybond-N nylon membrane (Amersham) in a Trans-Blot Electroblotting apparatus with plate electrodes (Bio-Rad). The appropriate region of the gel was lifted to blotting paper, immersed in 0.5x TBE and overlaid with Hybond-N membrane, avoiding air bubbles. The sandwich was placed between blotting paper and Scotchbrite pads, and placed in 0.5x TBE transfer buffer in the apparatus. Transfer was carried out for 30min at 60V, and RNA was crosslinked to the membrane by exposure to UV light in a UV Stratalinker 1800 (Stratgene).

2.4.11 Hybridisation of Southern and Northern Blots with Random-primed Probes.

Northern and Southern blots were hybridised to probes in sealed bottles on a rotisserie in a Hybaid hybridisation oven. To hybridise multiple blots to the same probe, up to three blots were placed in a single Hybaid bottle, interlaying with nylon mesh to prevent contact between membranes. The preparation of ^{32}P-labelled probes has been described (section 2.4.5).

1) Pre-hybridisation was carried out at 42°C for 2-4hr in 1x P buffer, 50% formamide, 1M NaCl, 100µg/ml denatured sonicated salmon sperm DNA (Sigma).

2) The probe was denatured by boiling for 3min and promptly cooled on ice.

3) The probe was added directly to the hybridisation bottle if less than 0.5ml, or if greater than 0.5ml mixed with 1 volume 4M NaCl and 2 volumes formamide, and then added.

4) Hybridisation was allowed to proceed for 14-18hr at 42°C.

5) The hybridisation solution was removed and stored at -20°C for reuse (probes were reused up to three times). The membrane was washed as follows:

(i) 2x10min at 42°C with 100ml 2x SSC.

(ii) 2x30min at 65°C with 100ml 2x SSC, 0.5% (w/v) SDS.

(iii) 2x30min at room temperature with 100ml 0.1x SSC.

6) Membranes were wrapped in saran wrap and exposed to X-ray film and/or phosphorimager screens.

7) To hybridise membranes to additional probes, blots were stripped by immersion in at least 2l of 0.1% (w/v) SDS at 100°C. Membranes were then treated as in the pre-hybridisation step (1).

Buffers:
5x P buffer:- 1% BSA, 1% polvinylpyrolidine (M.Wt. 40kD), 1% Ficoll (M.Wt. 400kD), 250mM Tris-HCl (pH7.5), 0.5% (w/v) sodium pyrophosphate, 5% (w/v) SDS.

2.4.12 Hybridisation of Northern Blots to Oligodeoxynucleotide Probes.

Northern blots of snRNA were hybridised to end-labelled oligdeoxynucleotides to probe for snRNAs by a modification of the method of Church and Gilbert (1984). Hybridisation was carried out in a Hybaid oven as above. The preparation of end-labelled probes has already been described (section 2.4.6).

1) Hybond-N membranes were prehybridised for at least 30min in 7% (w/v) SDS, 0.5M sodium phosphate (pH7.0), 1mM EDTA at 28°C.

2) The prehybridisation buffer was discarded and replaced with the same buffer containing ^{32}P-labelled oligodeoxynucleotides. Hybridisation was allowed to proceed for 14-18hr. The solution of labelled oligonucleotides (in 10mM sodium phosphate after gel filtration) was boiled for 3min before use.

3) The membrane was washed at least three times in 5% (w/v) SDS, 0.5M sodium phosphate (pH7.0), 1mM EDTA at 28°C, until the background radioactivity was negligible.

4) Membranes were blotted dry on filter paper to prevent diffusion of oligonucleotides from sites of hybridisation, wrapped in Saran Wrap and exposed to X-ray film and/or phosphorimager screens.

2.4.13 The Polymerase Chain Reaction (PCR).

PCR was used to identify recombinants, to generate DNA fragments for use as probes, to amplify regions of the yeast genome for cloning, as a means to create deletion mutations, and to generate sequencing templates. Each application followed the general protocol described below, or modifications upon it.

1) PCR reaction volumes (30-100µl) contained 1.5mM $MgCl_2$ (see below), 0.2mM each of dATP, dGTP, dTTP and dCTP, 0.25µM each of two oligodeoxynucleotide primers,

1x reaction buffer (supplied with enzyme), 10U/ml thermostable polymerase, and template DNA.

2) Reaction volumes were overlaid with mineral oil and subjected to temperature cycling in a Hybaid Thermal Reactor, programmed as follows:

(i) 93°C for 3min. 1 cycle

(ii) 92°C for 1min, 45°C for 1min, 72°C for 5min. 30 cycles.

(iii) 72°C for 10min, 45°C for 10min. 1 cycle.

3) 5-10µl reaction volume was fractionated on an agarose gel to assess yield of product. To separate products from oligonucleotide primers, dNTP and enzyme, Qiaquick Spin PCR Purification columns (Qiagen) were used according to the manufacturers instructions.

During the course of this research, thermostable polymerase was purchased from two sources. Initially, Thermalase (IBI) extracted from *Thermus thermophilus* was used. This enzyme was supplied with 10x reaction buffer (100mM Tris-HCl (pH8.3), 500mM KCl, 15mM $MgCl_2$, 0.01% Tween-20, 0.01% (w/v) Gelatin, 0.01% NP-40) containing a magnesium salt, and no additional $MgCl_2$ was added to the reaction. This product was discontinued. The second source was Advanced Biotechnologies Ltd, who supplied Taq polymerase, and Thermoprime^{Plus} polymerase extracted from *Thermus icelandicus*. These enzymes were supplied in 50% (v/v) glycerol, 10mM potassium phosphate (pH7.0), 100mM NaCl, 0.1mM EDTA, 2mM DTT, and were diluted in the same buffer. The 10x reaction buffer supplied with these enzymes was "Buffer IV" (200mM ammonium sulphate, 750mM Tris-HCl (pH9.0), 0.1% (w/v) Tween) which did not contain a magnesium salt, and $MgCl_2$ was added to the reaction.

2.4.14 *In vitro* Transcription.

In vitro transcription reactions were performed as described in Davanloo *et al.* (1984), with phage T7 RNA polymerase.

1) 4µg of template plasmid (pT7rp28) was linearised by incubation with 10U *Eco*RI in a 20µl reaction volume (1x buffer H) at 37°C for 2hr. Linearised template was stored at -20°C.

2) 40μCi [α-^{32}P]UTP (800Ci/mmol) was dried down *in vacuo* and resuspended in 2.2μl dH_2O.

3) The following reaction mix (9.5μl) was assembled: 2μl restriction digest (from step 1), 1μl each of ATP, CTP, GTP at 5mM, 1μl 250μM UTP, 1μl 10x T7 buffer, radiolabel in 2.2μl, 0.3μl RNasin (10U; Promega).

4) 0.5μl T7 RNA polymerase was added and the reaction mix incubated at 37°C for 12min.

5) Reactions were stopped on ice.

6) 0.1-0.5μl of the reaction mix was subjected to electrophoresis on a small sequencing gel (16cm x 16cm plates) to determine product purity and yield.

7) Aliquots of the reaction volume were used directly in splicing reactions, or stored at -20°C for up to 3 days.

8) Alternatively, transcripts were separated from unincorporated nucleotide by adding to the reaction mix 20μl 7.5M ammonium acetate, 1μl 20mg/ml *E.coli* tRNA, 100μl dH2O, and 400μl ice-cold ethanol. After incubation for 10min on dry ice, precipitated RNA was sedimented in a microfuge (4°C, 10min), and the supernate discarded. The tube was respun, and the residual supernate discarded.

9) The pellet was allowed to dry in air for 10min, redissolved in 10μl dH_2O, and stored as above.

10x T7 buffer:- 400mM Tris-HCl (pH8.0), 100mM $MgCl_2$, 100mM DTT, 100mM NaCl.

2.4.15 Purification of Large Restriction Fragments for Cloning.

The Quiex kit (Quiagen) for DNA fragment purification by absorption to glass beads in the presence of a chaotropic salt was used to purify fragments less than 5kb in length. However, it was found to cause degradation of larger fragments (probably due to shearing) and to inhibit ligation. For use in cloning, larger fragments were purified by removing agarose from a molten gel slice by phenol extraction, as below. This method enabled the recovery of large restriction fragments in a high-integrity form which were efficient substrates of ligation.

1) The gel slice (not more than 0.1ml volume) containing the appropriate restriction fragment was excised from low melting point agarose gel and weighed.

2) 4 volumes of TE (20mM Tris-HCl (pH8.0), 1mM EDTA) was added, and incubated at 65°C for 5min to melt agarose.

3) One phenol extraction was carried out and the supernate transferred to a fresh microfuge tube, avoiding the agarose at the interface.

4) After further phenol/chloroform extraction (until interface was clear), $1/10^{th}$ volume 3M sodium acetate (pH5.3), 1µl 20mg/ml glycogen as a carrier (Boehringer, molecular biology grade) and 2 volumes of ethanol were added. DNA was precipitated at -20°C for 20min.

6) The pellet was washed with ice-cold 70% (v/v) ethanol, dried *in vacuo*, and dissolved in dH_2O (10-20µl). Recovery was confirmed by agarose gel electrophoresis.

This method routinely recovered at least 50% of the restriction fragment.

2.5 PROTEIN AND IMMUNOLOGICAL METHODS.

2.5.1 Quantitation of Proteins.

Protein concentration was estimated using the Bradford assay (Bradford, 1976) with a dye concentrate purchased from Bio-Rad and used according to the manufacturers instructions. Standard curves were generated using BSA.

2.5.2 SDS-Polyacrylamide Gel Elecrophoresis (SDS-PAGE) of Proteins.

Proteins were separated on discontinuous polyacylamide gels containing SDS as described by Laemmli (1970). 6% gels were cast between 16cm x 16cm glass plates with 1.5mm spacers and 0.5cm wells. The stacking gel was the same depth as the wells. Samples were denatured by adding an equal volume of protein loading buffer, heating to 100°C for 5min, and centrifuging for 3min to remove any precipitated material. After denaturation, samples were loaded and run within 15min to avoid degradation. Gels were run in an ATTO electrophoresis tank at 30-150V in SDS-PAGE buffer.

Buffers:

Protein loading buffer:- 125mM Tris-HCl (pH6.8), 200mM DTT, 4% (w/v) SDS, 40% (v/v) glycerol, 0.02% (w/v) bromophenol blue. Stored at 4°C for not more than 1 month.
SDS-PAGE buffer:- 25mM Tris, 192mM glycine, 0.1% (w/v) SDS.

2.5.3 Coomassie Staining of Protein Gels.

SDS-PAGE gels were stained to visualise proteins by submersing in Coomassie staining solution (25% (v/v) methanol, 10% (v/v) acetic acid, 0.05% (w/v) Coomassie Brilliant Blue G) and incubating at 37°C for 60min. Gels were destained with 10% (v/v) acetic acid, 10% (v/v) methanol at 37°C until the background was clear.

2.5.4 Affinity Purification of Antibodies.

Antibodies were affinity purified by the method of Robinson *et al.*, (1988) by binding to the peptide or fusion protein against which the antibody was raised. Strips of nitrocellulose onto which the fusion protein (or a peptide-BSA conjugate in the case of anti-8.6 antibodies) had been blotted were prepared previously (Anderson, 1989; Brown, 1992).

1) Strips were incubated with blocking buffer (1% (w/v) ovalbumin, 1x TBS, 0.1% (w/v) sodium azide) for 2-4hr at 37°C, or overnight at room temperature.

2) The buffer was discarded and replaced with 10ml of the same buffer containing 30-100µl crude polyclonal serum, and incubated for 2-3hr at room temperature or overnight at 4°C.

3) Strips were washed in 1x TBS, 0.1% NP-40 (5 washes, 12min each wash).

4) To elute bound antibodies, strips were incubated with 2ml 0.1M glycine-HCl (pH2.5), 1% (w/v) ovalbumin for 5min, and then with 2ml same buffer for 10min, both at room temperature. Each elution volume was neutralised immediately upon removal with 0.5ml 1M Tris-HCl (pH7.5).

5) Eluate fractions were pooled and dialysed against 3l of 1x TBS with 3 changes of buffer, for 16hr at 4°C.

6) Sodium azide was added to 0.1% (w/v), and aliquots were stored at -20°C.

2.5.5 Western Blotting.

Following fractionation by SDS-PAGE, proteins were transferred by electrophoresis to Immobilon-P membrane (Millipore) in a Trans-Blot Electroblotting Cell (Bio-Rad) with plate electrodes. The membrane was reversibly stained with Ponceau S (Sigma) to verify transfer, and incubated with anti-Prp8p antibodies.

1) Immobilon-P membrane was cut to the size of the gel, and activated by rinsing in methanol.

2) The gel was assembled with the Immobilon-P membrane between blotting paper and Scotchbrite pads following the manufacturers instructions.

3) The assembly was placed in transfer buffer in the electroblot apparatus with two cooling coils and electrophoresis run for 3hr at 100V, or at 60V overnight.

4) After transfer the membrane was stained with Ponceau S, and molecular weight markers cut off. Ponceau S was removed by rinsing with water.

5) The membrane was incubated in blocking buffer for 1hr.

6) Blocking solution was discarded and replaced with the anti-Prp8p antibody (1:1000 dilution of crude serum or 1:20 dilution of affinity purified antibody in blocking solution), and incubation continued for 1-2hr at room temperature or overnight at 4°C.

7) The blot was washed four times in 1x TBS, 15min each.

8) The blot was then incubated with the secondary antibody (goat anti-rabbit antibodies, coupled to alkaline phosphatase (Promega) at a dilution of 1:7500 in 15ml 1x TBS) for 1hr.

9) as 7)

10) The blot was rinsed in dH_2O and incubated with development solution until bands were visible.

11) The development solution was discarded and the blot rinsed with dH^2O and blotted dry. The blot was wrapped in Saran Wrap and stored in the dark to prevent fading.

Buffers:

10x Ponceau S:- 2% (w/v) in 30% (w/v) trichloroacetic acid and 30% (w/v) sulphosalicylic acid.

Transfer buffer:- 20mM Tris, 150mM glycine.

Blocking buffer:- 5% (w/v) BSA, 1x TBS.

Development solution:- 100mM Tris-HCl (pH9.5), 100mM NaCl, 50mM $MgCl_2$, 0.33mg/ml NBT (Promega), 0.165mg/ml BCIP (Promega).

2.5.6 Immunoprecipitation.

Splicing complexes containing Prp8p were immunoprecipitated from mock splicing reactions (splicing reactions with no substrate pre-mRNA) by a modification of the method of Lossky *et al.*, (1987). Antibodies were pre-bound to protein A conjugated to Sepharose beads (PAS; Sigma).

1) PAS beads (12.5mg per sample) were swollen in NTN for 25min and washed 5 times with 1ml NTN.

2) Antibodies were pre-bound to PAS by adding 15µl antiserum and 0.1ml NTN for every 12.5mg PAS (original weight) and incubating for 1hr at room temperature on a turning wheel.

3) PAS-coupled antibodies were washed 5 times with 1ml NTN, and stored for up to 36hr at 4°C.

4) Mock splicing reactions were mixed with an equal volume of IP buffer and added to the PAS-coupled antibodies.

5) The mix was incubated for 2-4hr on a turning wheel at 4°C.

6) Beads were sedimented and the supernate transferred to a fresh tube and retained. Beads were washed four times with 1ml NTN and then once with 1ml NT, removing the last wash completely.

7) To analyse proteins, PAS beads were boiled with protein loading buffer and subjected to SDS-PAGE (section 2.5.2). To analyse snRNAs, the procedure below was followed:

8) 100µl proteinase K solution was added and incubated with agitation for 30min at 37°C.

9) 50µl dH_2O was added and a phenol/chloroform extraction performed. The PAS/phenol phase was back-extracted to a further 50µl dH_2O and the aqueous phases pooled.

10) Two further phenol/chloroform extractions were performed, and RNA precipitated by addition of 2µl 20mg/ml tRNA and 3 volumes ice-cold ethanol, and incubation on dry-ice for 20min.

11) RNA was sedimented and the pellet resuspended in 10µl dH_2O.

12) 7µl Sanger dye was added and the sample denatured (95°C, 3min) and fractionated through a sequencing gel. SnRNAs were detected by Northern blotting (2.4.12).

An aliquot of the immunoprecipitation supernate (from step 6) and/or aliquots of neat splicing extract were treated with proteinase K and analysed for snRNA content (step 8 and onwards) to assess the efficiency of immunoprecipitation.

Buffers:
NTN: 50mM Tris-HCl (pH7.5), 150mM NaCl, 0.1% (v/v) NP-40.
NT: as NTN except without NP-40.
IP buffer: 10mM HEPES-KOH (pH7.5), 192.5mM KPO_4, 2.5mM $MgCl_2$, 0.2% (v/v) NP-40.
Proteinase K solution: 50mM Tris-HCl (pH7.5), 300mM NaCl, 5mM EDTA, 1.5% (w/v) SDS, 2mg/ml proteinase K (Beohringer). Store at -20°C.
Sanger dye: 50% (v/v) formamide, 20mM EDTA, 0.3% xylene cyanol FF, 0.3% bromophenol blue.

2.5.7 *In vitro* Splicing.

2.5.7.1 Preparation of Yeast Splicing Extract.

Splicing extracts were prepared by a modification of the method of Lin *et al.*, (1985).

1) 2-4l of an appropriate medium was inoculated from a fresh stationary culture and grown overnight to OD_{600} 0.5-1.0.

2) Cells were harvested, washed in 50ml 50mM potassium phosphate (pH7.5), and resuspended in 40ml lyticase buffer.

3) To prepare sphaeroplasts by digestion of the cell wall, 3700U of lyticase was added in 1-2ml lyticase buffer. Cells were incubated at 30°C, or at 23°C for temperature-sensitive strains, with gentle agitation to maintain suspension.

4) Digestion was allowed to proceed for 40-60mins, until most cells had formed sphaeroplasts (assessed by sensitivity of samples to lysis in 0.1% (w/v) SDS).

5) Sphaeroplasts were transferred to a 50ml Greiner tube, spun down (3k rpm, 5min), and washed twice with 1.2M sorbitol. Sphaeroplasts were resuspended at each step with a glass rod.

6) (Can be omitted) To heat-inactivate a temperature-sensitive strain, sphaeroplasts were resuspended in 200ml YPDAS and transferred to a 500ml flask. Sphaeroplasts were incubated for 2hr at 36°C (or 37°C for *dbf3-1* strains) with gentle agitation, and harvested (3k rpm, 5min).

7) The pellet was washed once in ice-cold SB-3 buffer, and harvested (3k rpm, 5min). Henceforth all procedures were carried out at 4°C.

8) The pellet was weighed, resuspended in ice-cold buffer A (1ml per gram pellet), and transferred to a chilled Dounce homogeniser and placed on ice for 5min.

9) The suspension was homogenised with 10-13 gentle strokes, transferred to a beaker on ice, and stirred with a magnetic bar. $1/9^{th}$ volume 2M KCl was added dropwise, and stirring continued for 30min.

10) The lysate was transferred to a cold, rinsed polycarbonate SS-34 tube and spun at 17k rpm for 30mins.

11) The supernate was removed, avoiding the top (lipid) layer as much as possible, and transferred to an ultracentrifuge tube and spun at 37k rpm for 1hr at 4°C in a pre-cooled Ti50 rotor.

12) The supernate was transferred to dialysis tubing, again avoiding the lipid layer, and dialysed for 3hr against buffer D, with three changes of buffer.

13) The extract was transferred to microfuge tubes and spun for 10min (4°C) to remove precipitated material. Aliquots were snap-frozen and stored at -70°C. Samples were not thawed and refrozen more than three times, as loss of activity occurs.

Buffers:

Lyticase buffer:- 1.2M sorbitol, 50mM potassium phosphate (pH7.5), 30mM DTT.

YPDAS:- YPDA with additionally 1.2M sorbitol.

SB-3:- 50mM Tris-HCl (pH7.5), 10mM $MgCl_2$, 1.2M sorbitol, 3mM DTT.

Buffer A:- 10mM HEPES-KOH (pH7.5), 1.5mM $MgCl_2$, 10mM KCl, 0.5mM DTT.

Buffer D:- 20mM HEPES-KOH (pH7.5), 0.2mM EDTA, 50mM KCl, 20% (v/v) glycerol, 0.5mM DTT.

DTT was added to solutions immediately prior to use.

2.5.7.2 The *In vitro* Splicing Reaction.

Splicing reactions were performed essentially as described in Lin *et al.*, (1985). The standard reaction of 10µl was scaled up two-fold to analyse both RNA products by denaturing electrophoresis and splicing complexes by native gel electrophoresis. Radiolabelled substrates were produced by *in vitro* transcription.

1) The following reaction mix was assembled: 50µl 30% (w/v) PEG-8000, 40µl IgG buffer, 30µl 1M KPO_4 (pH7.5), 12.5µl 100mM $MgCl_2$, 10µl 100mM ATP, approximately 2.5µCi radiolabelled rp28 transcript, made up to 250µl with sterile dH_2O.

2) 10µl aliquots of splicing mix were transferred to sterile microfuge tubes and 10µl splicing extract added.

3) Reactions were incubated for 25min in water bath and quenched on ice.

4) To 10µl reaction volume, 2µl proteinase K solution was added and thoroughly mixed, and incubated at 37°C for 30min.

5) 100µl 'splicing cocktail_ buffer was added, and phenol/chloroform extracted twice.

6) RNA was precipitated with 500µl ice-cold ethanol and 2µl 20mg/ml *E.coli* tRNA as carrier, and after centrifugation the pellet resuspended in 3µl Sanger dye.

7) Samples were denatured (3min, 95°C), fractionated through sequencing gels and exposed to autoradiography.

Buffers:

'splicing cocktail_:- 50mM sodium acetate (pH5.3), 1mM EDTA, 0.1% (w/v) SDS, 25µg/ml *E.coli* tRNA.

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

IgG buffer:- 0.5M KPO_4 (pH7.5), 33mM citrate (pH3.0), 16.67% (v/v) glycerol.

2.5.8 Native Gel Electrophoresis of Splicing Complexes.

Native (non-denaturing) gel electrophoresis was carried out as described in Pikielny *et al.* (1986) except that the EDTA concentration of the gel was increased from 1mM to 10mM. This composite agarose/polyacrylamide gel system resolves three complexes designated I, II, and III, in order of increasing electrophoretic mobility. Complex II, the last to form, contains the RNA intermediates and products of splicing (Brown, 1992), and is therefore regarded as the active spliceosome.

1) Agarose (0.25g) was dissolved in 40ml dH_2O by heating and cooled to about 55°C.

2) 5ml 30% (w/v) acrylamide (60:1 acrylamide:bisacrylamide), 2.5ml 10x TB, and 1ml 0.5M EDTA were pre-heated to 65°C and added to the agarose.

3) 0.4ml 10% (w/v) ammonium persulphate and 50µl TEMED were added, the volume made up to 50ml (dH_2O), and mixed by inversion.

4) The gel was poured between prewarmed 16cm x 16cm plates with a 1.5cm spacer and 10 x 9mm slot comb, and allowed to set at 37°C for 30min, and then 4°C for at least 30min. The gel was pre-run in an ATTO electrophoresis tank in 0.5x TB, 10mM EDTA at 4°C for 1hr.

5) 10µl splicing reaction (after stopping on ice, step 3, section 2.5.7 above) was mixed with 10µl buffer Q and incubated on ice for 10 min.

6) 6µl RNP loading buffer was added, and the samples fractionated for 6hr at 80V or overnight at 40-50V.

7) The gel was wrapped in Saran Wrap and exposed to autoradiography without drying.

Buffers:

Q buffer: 450mM KCl, 2.3mM magnesium acetate, 23mM EDTA, 18mM Tris-HCl (pH7.5), 1.2mg/ml yeast total RNA (prepared as in section 2.4.4).

RNP loading buffer: 5x TB, 50mM EDTA, 50% (v/v) glycerol, 0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF.

Chapter 3.

Construction of *PRP8* Expression Systems and Truncation Mutants.

3.1 Introduction.

Functional domains within proteins are broadly divisible into two types: those which are constituted by discontinuous peptide regions dispersed throughout the primary sequence of a protein which unite during tertiary structure formation into structural motifs, and those which are constituted by peptide regions contiguous in the primary sequence. The most familiar examples of the latter type are from the family of transcription regulatory factors: in many cases, after removal of a proportion of the protein, the remnant fragment possesses DNA binding activity similar to the full-length protein (eg Gal4p). Also, the peptide regions of these factors which influence transcription can be fused to heterologous DNA-binding domains with retention of activity. Hence domains of the second type may be readily delineated, and their functions characterised, by deletion mutagenesis. This can be a useful starting point for detailed studies of structure, as functional regions and structural domains are likely to coincide. Analyses of this type can also suggest how a protein might have evolved, by the combination of functional modular domains.

Within the Prp8p splicing factor, the proline-rich acidic N-terminal domain (Figure 3.1) is an obvious candidate for deletion mutagenesis. From consideration of the physical properties of proline (Section 1.11.2), this region is expected to form a discrete structural domain. Some evidence for this has been furnished from experimentation: antibodies directed specifically against this region (anti-8.6 antibodies) recognise the native protein in various spliceosomal complexes, suggesting that the N-terminal domain is exposed to solution (at least partially) at multiple stages of the splicing pathway (Anderson, 1989; Brown, 1992). In other proteins, proline-rich regions have been shown to participate in protein-protein interaction and in processes to assemble macromolecular complexes (Williamson, 1994; also Section 1.12). Therefore,

a reasonable expectation was that a truncated version of Prp8p lacking the N-terminal domain would be defective in one particular aspect, but that other functions of this multifunctional factor would be relatively unaffected.

When the N-terminal domain is deleted, the truncated allele does not complement a *prp8*-null strain (J.D. Brown, unpublished results). This was encouraging, as it suggested that the domain had some essentail function. The next objective was to generate a mutant with a conditional phenotype which would be amenable to further analysis. This chapter describes the placement of the truncated allele downstream of the *GAL1* promoter for overexpression, and the construction of mutants truncated of part of the N-terminal domain. For ease of reference, plasmids described here and in Chapter 4 and presented in Table 3.1 (see also Appendix A).

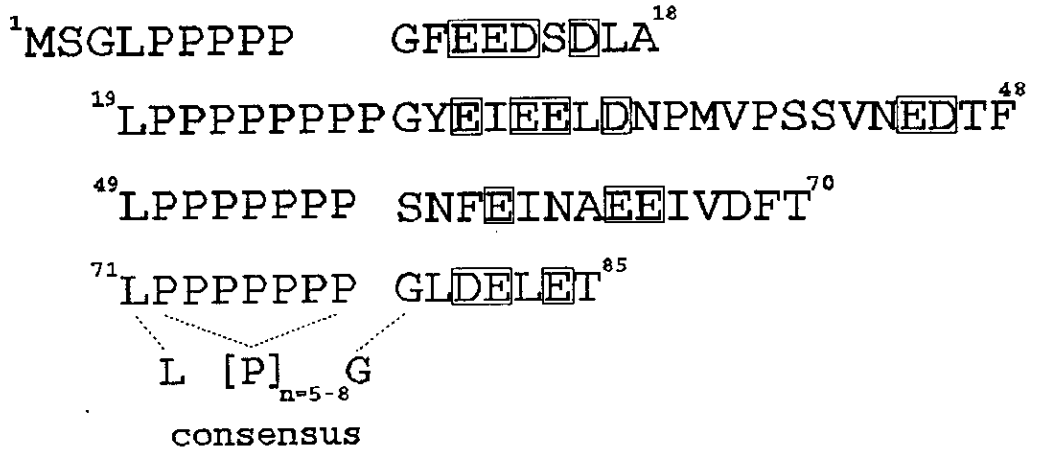


Figure 3.1 Amino Acid Sequence of N-Terminal Domain of Chromosomal *PRP8*. The strings of contiguous prolines are aligned; acidic residues are boxed. Introduction of the *Nhe*I site changed position 5 (P to A).

Notes:

(1): "JDY" plasmids were constructed by J.D.Brown.

(2): As described by Brown & Beggs (1992), *PRP8* was mutagenised at two nucleotide positions to introduce an *Nhe*I site adjacent to the start codon. This changed the amino acid at position 5 (P to A; see Figure 3.1). pY8000 and pY8500 bear *PRP8* with no *NheI* site; in all other (later) plasmids in the table *PRP8* contains the *Nhe*I site. For simplicity, and as this mutagenesis has not been detected to affect the activity of the protein (Brown and Beggs, 1992), both versions of *PRP8* (with and without *Nhe*I) are regarded as wild-type in this thesis.

(3): Deletions (Δ) are defined by amino acid number.

Table 3.1: Table of Plasmids.

(A): Vectors

Name	Derived from	Features
pRS316	pBluescript	*CEN, ARS, URA3* Multiple Cloning Site
pBM125	YCp50	*CEN, ARS, URA3*
YEp24	-	2μ, *URA3*

(B): *PRP8* alleles

Name[1]	Features	Derived from	Allele[2,3]	Copies of LP_nG	Promoter
pY8000	2μ, *URA3, PRP8*	YEp24	wild-type	4	*PRP8*
pY8500	*CEN, ARS, URA3, PRP8*	YCp50	wild-type	4	*PRP8*
pJDY6	2μ, *URA3, PRP8*	YEp24	wild-type	4	*PRP8*
pAJY12	*CEN, ARS, URA3, PRP8*	pRS316	wild-type	4	*PRP8*
pJDY10	*CEN, ARS, URA3, GAL1UAS:PRP8*	pRS316	wild-type	4	hybrid *GAL1UAS:PRP8*
pJDY13	*CEN, ARS, HIS3, GAL1:PRP8*	pRS313	wild-type	4	*GAL1* (glucose repressed, weak expression)
pBM-PRP8	*CEN, ARS, URA3, GAL1:PRP8*	pBM125	wild-type	4	*GAL1-10* (overexpression)
pAJY11	*CEN, ARS, URA3, GAL1:PRP8*	pRS316	wild-type	4	*GAL1* (overexpression)
pAJY13	*CEN, ARS, URA3, prp8ΔP2*	pRS316	Δ29 to 79	2	*PRP8*
pAJY1	*CEN, ARS, URA3, GAL1:prp8ΔP2*	YEp24	Δ29 to 79	2	*GAL1* (overexpression)
pJDY7	2μ, *URA3, prp8ΔN*	YEp24	Δ4 to 79	none	*PRP8*
pAJY9	*CEN, ARS, URA3, GAL1:prp8ΔN*	pBM125	Δ4 to 79	none	*GAL1* (overexpression)

3.2 PLASMIDS BEARING WILD-TYPE *PRP8*.

3.2.1 A System to Conditionally Overexpress *PRP8*.

3.2.1.1 Construction.

Overexpression in yeast has been achieved by fusion to several different regulated promoters, including those of the *GAL1*, *PHO5*, *ADH1* and *PGK* genes (Guthrie & Fink, 1991). The *GAL1* promoter in particular is well characterised and provides regulation of expression over a wide range dependent on growth conditions (for a review of regulation of galactose metabolism (*GAL*) genes see Johnston, 1987). Transcription from the *GAL1* promoter is strongly induced by galactose, and strongly repressed by glucose, characteristics which are suitable for the propagation of dominant negative mutants (Liu *et al.*, 1992), and therefore of importance when creating mutants with unknown phenotypes.

At the outset of this project, two *PRP8* expression systems based on the *GAL1* promoter were available (Brown, 1992; Brown and Beggs, 1992). In pJDY10 (*CEN, URA3, GAL1-10 UAS:PRP8*), *PRP8* is transcribed from a hybrid promoter consisting of the upstream activation sequence (UAS) of the *GAL1* promoter and the region upstream of *PRP8*. In pJDY13 (*CEN, HIS3, GAL1:PRP8*) the *PRP8* open reading frame is fused to the entire *GAL1* promoter. These plasmids are not ideal for determining the phenotype of mutants when overexpressed because neither enables regulated overexpression. pJDY10 expresses *PRP8* even on glucose and repression is inefficient (the extent of galactose induction with pJDY10 is not known). pJDY13 is glucose-repressible, but produces less than wild-type levels of Prp8p under inducing conditions (galactose as carbon source), probably because the context of the ATG start codon is suboptimal for translational initiation (Brown, 1992).

To improve production of Prp8 protein, pBM-PRP8 (*CEN, URA3, GAL1:PRP8*) was created. This is another fusion of the *GAL1* promoter to the *PRP8* open reading frame, but with additionally a linker encoding sequence around the ATG start codon to match the consensus for strongly-expressed yeast genes (Hamilton *et al.*, 1987). Construction of pBM-PRP8 is detailed in Figure 3.2.

Figure 3.2A: Construction of pBM-PRP8.

Plasmid pBM-PRP8 (*CEN, URA3, GAL1:PRP8*) was constructed in three steps:

Step 1: The *Nhe*I-*Sal*I region of *PRP8* and a *Bam*HI-*Nhe*I linker were inserted between the *Bam*HI and *Sal*I sites of pBM125 (*CEN, URA3*), in a three-fragment ligation reaction. pBM125 is a derivative of YCp50 containing the 0.8 kb region between the divergently transcribed *GAL1* and *GAL10* genes (*GAL1-10*). The linker consisted of oligonucleotides 976N and 977N. It encodes the first 3 codons of *PRP8,* and contains a *Bss*HII site and the sequence AACACA preceding the ATG start codon. The second codon of the chromosomal *PRP8* allele is AGT, encoding serine. In the 976N/977N linker, this is changed to TCT (also encoding serine). The context of the *PRP8* start codon in pBM-PRP8 conforms to the consensus for strongly expressed yeast genes [(A/T)A(A/C)A(A/C)A**ATG**TC(T/C), start codon in bold; Hamilton *et al.*, 1987]

Step 2: The three-fragment ligation product from step 1 was cleaved and religated at the unique *Bss*HII site in the linker, to produce pAJY6. This ensured monomeric linker insertion.

Step 3: The remainder of *PRP8* was inserted as a *Sal*I-*Eag*I fragment.

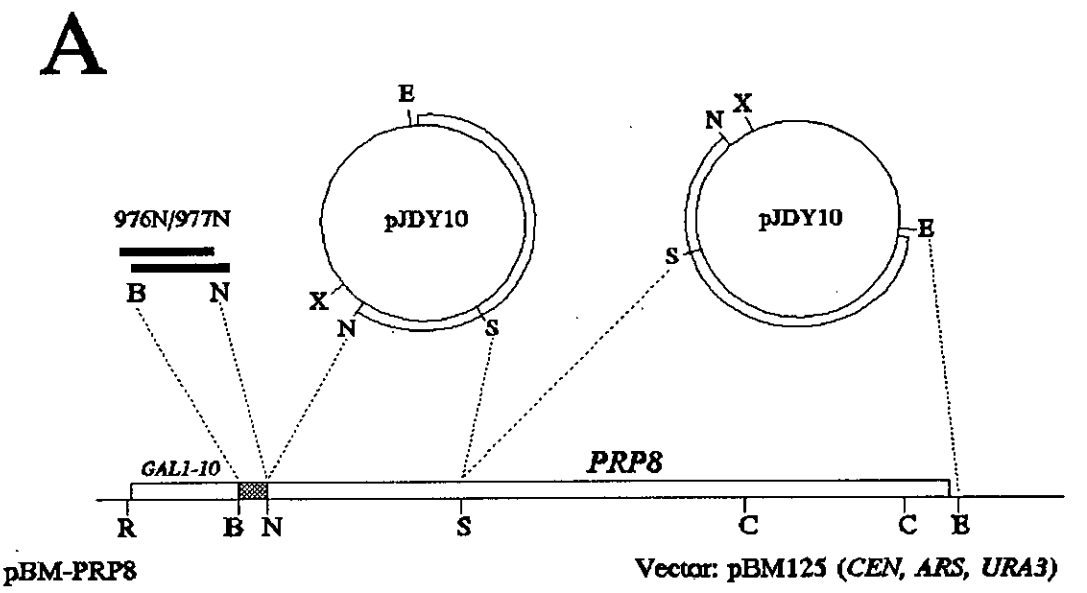
Figure 3.2B: Construction of pAJY12.

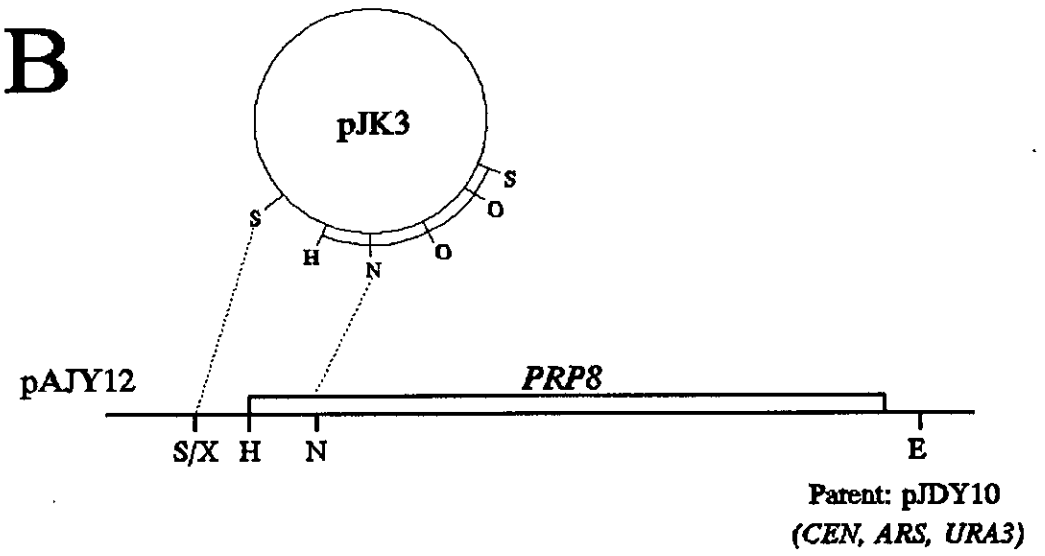
The transcription control region of the *PRP8* locus was removed from pJK3. pJK3 is the plasmid upon which the mutagenesis to create the *Nhe*I site was carried out (by J.Keddy; Brown and Beggs, 1992). It consists of the *Sal*I fragment from pY8500 (encoding 1.8kb of *PRP8* coding sequence and the upstream region) inserted into pBR322.

To create pAJY12, the promoter region of pJDY10 was removed as an *Xho*I-*Nhe*I fragment and replaced with the *Sal*I-*Nhe*I fragment from pJK3 containing the *PRP8* upstream region and the first 3 codons of *PRP8*. The *Nhe*I site is unique in pAJY12, enabling the placement of mutant *PRP8* alleles under the transcriptional control of the *PRP8* promoter.

Key.

B: *Bam*HI, N: *Nhe*I, X: *Xho*I, S: *Sal*I, E: *Eag*I, C: *Cla*I, O: *Nco*I, H: *Hind*III, S/X: fusion of *Sal*I and *Xho*I complementary ends.


A
E
pJDY10
976N/977N
B
N
X
N
S
N
X
pJDY10
S
E
GAL1-10
PRP8
R
B
N
S
C
C
B
pBM-PRP8
Vector: pBM125 (CEN, ARS, URA3)


B
pJK3
S
O
S
H
N
O
pAJY12
PRP8
S/X
H
N
E
Parent: pJDY10
(CEN, ARS, URA3)

A second and similar construct was pAJY11 (*CEN, URA3, GAL1:PRP8*) which was created for use as a control in several experiments, and to maximise the utility of the *GAL1:PRP8* system in cloning. The *CEN* and *URA3* elements of pAJY11 derive from a different vector to pBM-PRP8, also pAJY11 does not contain the entire *GAL1-10* region, and its promoter was created by PCR. Otherwise the organisation of promoter fused to *PRP8* coding sequence (with optimal translational start) is identical to pBM-PRP8. Construction of pAJY11 is detailed in Figure 3.7.

3.2.1.2 Characterisation.

The plasmids pAJY11 and pBM-PRP8 (both *GAL1:PRP8*) are expected to be functionally equivalent. They were introduced into four genetic backgrounds (wild-type, *prp8* mutant, *ded1* mutant, and *prp8* null) and tested by both genetic and biochemical means to confirm that they exhibited the expected pattern of expression.

1) Temperature-sensitive backgrounds: Cells of the temperature-sensitive *prp8-1* strain SPJ8.31 bearing plasmid pBM-PRP8 (*CEN, URA3, GAL1:PRP8*) or pBM125 (*CEN, URA3*) were suspended in sterile distilled water and spotted for growth under various conditions, maintaining selection for the *URA3* plasmid. On YMGRcas plates the control (pBM125) transformants were temperature-sensitive above 30°C, as expected for this strain. Under the same (inducing) conditions, pBM-PRP8 confers the ability to grow at 34°C (Figure 3.3A) and 36°C (not shown), indicating that the plasmid produces functional Prp8p. On YMGlucas plates (repressing conditions), SPJ8.31 bearing pBM-PRP8 exhibits weak growth at 34°C, significantly more than with the negative control (pBM125), but far less than with pJDY6 (2μ, *URA3, PRP8*; not shown). This demonstrates that pBM-PRP8 is glucose-repressible, and that repression reduces *PRP8* expression to levels limiting for growth. As weak growth can be detected, repression is not complete, and *GAL1:PRP8* is leaky. On glucose, the pBM-PRP8 transformants had a high frequency of apparent reversion to a non-temperature-sensitive state. As this was specific for the *GAL1:PRP8* plasmid, it was probably caused by mutation relaxing the extent of glucose repression. This resulted in strong growth of a small proportion of cells, distinct from weak complementation which causes slow growth of all cells.

2) *ded1* background: The strain DJY105 contains *spp81-3*, a mutant allele of the *DED1* gene. This and other *ded1* alleles were isolated as suppressors of the *prp8-1* mutation (Jamieson *et al.*, 1991). The *spp81-3* point mutation introduces a stop codon into the *DED1* open reading frame. Subsequent work has shown that reduced expression of wild-type *DED1* also suppresses *prp8-1* (S. Maccallum and J.D.Beggs, unpublished data). This is believed to be the basis of suppression; as the *DED1* gene is essential, *spp81* strains must contain a low level of *DED1* activity due to translational readthrough. Jamieson *et al.* (1991) showed that overexpression of wild-type *PRP8* was toxic in *spp81* strains. Cells of the *spp81-2* genotype bearing pY8500 (*CEN, URA3, PRP8*) grew poorly, and with the high copy-number plasmid pY8000 (2μ, *URA3, PRP8*), *spp81-2* transformants could not be obtained (Jamieson *et al.*, 1991). To characterise the Prp8 protein level due to *GAL1:PRP8*, pAJY11 (*CEN, URA3, GAL1:PRP8*) and pRS316 (*CEN, URA3*) were introduced into DJY105 (*spp81-3*). Ura^+ transformants were propagated under repressing conditions (YMGlucas plates). On glucose, DJY105 bearing pAJY11 grew slower than the same strain bearing pRS316 (Figure 3.3B) confirming that *GAL1:PRP8* expresses the gene at low levels under repressing conditions. On galactose, DJY105 bearing pAJY11 grew very poorly indeed (Figure 3.2B). Thus toxicity due to *GAL1:PRP8* expression was at least as severe as that with pY8500 reported by Jamieson *et al.* (1991). This genetic evidence suggests that when induced, *GAL1:PRP8* produces levels of Prp8 protein at least as great as the chromosomal locus.

3) The null background: Haploid strain JDY8.05 (*ura3, his3, prp8::LEU2,* pJDY13) contains a null allele at the chromosomal locus: *PRP8* is deleted and replaced with *LEU2*. The glucose-repressible plasmid pJDY13 (*CEN, HIS3, GAL1:PRP8*) provides functional *PRP8*, supporting growth on galactose. The strain is not viable on glucose (Brown and Beggs, 1992). *URA3* plasmids allow pJDY13 to be lost from this strain by plasmid shuffle, if they confer *PRP8* function. To test pBM-PRP8 (*URA3*) for *PRP8* function, JDY8.05 was transformed with the plasmid, isolating colonies bearing both pBM-PRP8 and pJDY13 on YMGR lacking both uracil and histidine. After single

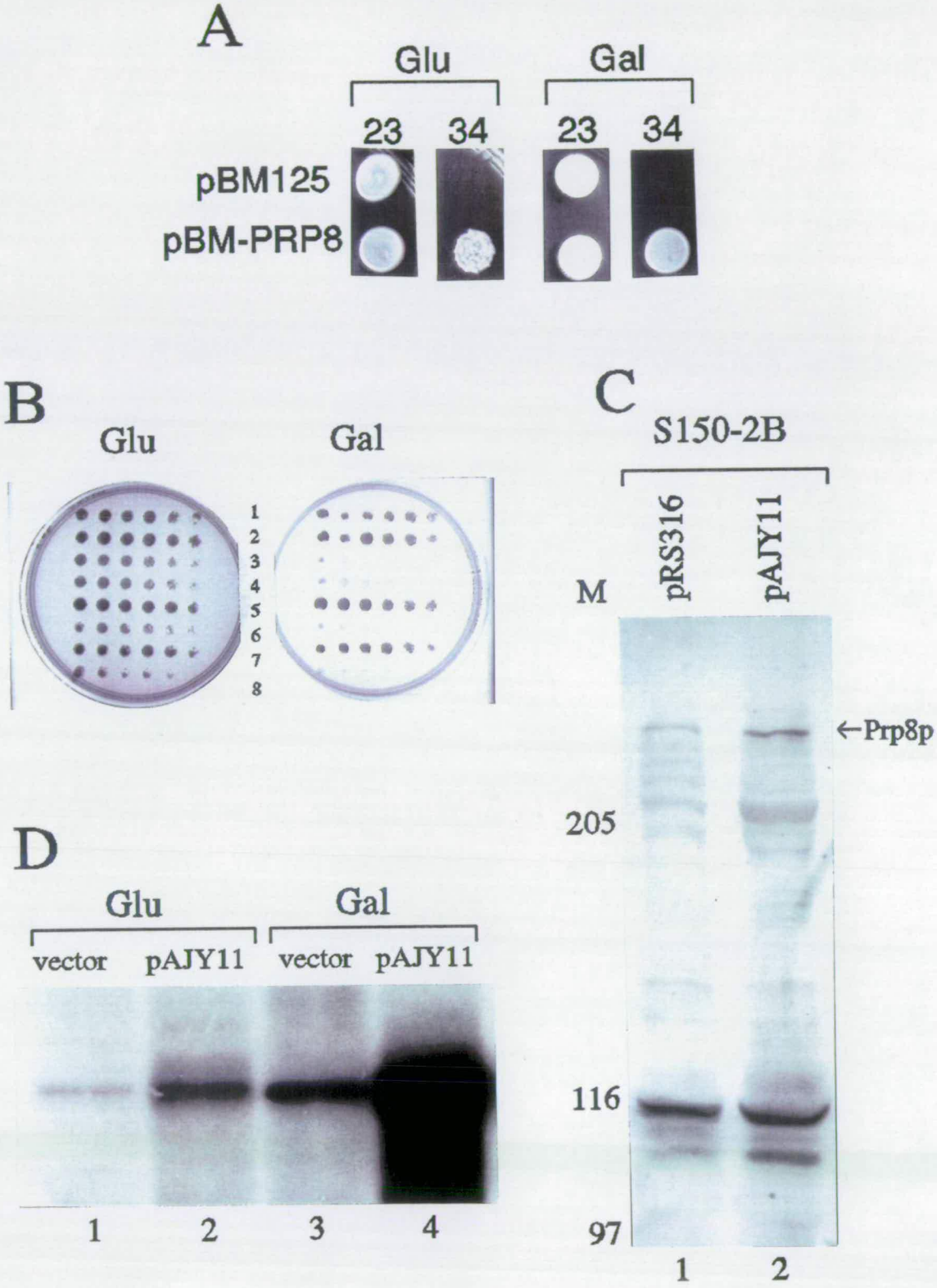
Figure 3.3: Characterisation of *GAL1:PRP8* Constructs.

(A): Complementation of SPJ8.31 (*prp8-1*). SPJ8.31 (*prp8-1*) cells bearing pBM-PRP8 (*CEN, URA3, GAL1:PRP8*) or pBM125 (*CEN, URA3*) were suspended in sterile distilled water at approximately 10^6 cells/ml, and transferred as 10µl droplets to YMGlucas (repressing conditions) and YMGRcas (inducing conditions) plates. Plates were incubated under permissive (23°C) or restrictive (34°C) conditions for 5 days.

(B): Lethality in DJY105 (*spp81-3*). DJY105 is a strain bearing a mutation in the *DED1* gene; overproduction of Prp8p in this strain is toxic. Plasmid pAJY11 (*CEN, URA3, GAL1:PRP8*) is derived from pBM-PRP8 and expected to be functionally equivalent (see main text). DJY105 cells bearing pAJY11 (rows 3,4,6,8) or pRS316 (*CEN, URA3*; rows 1,2,5,7) were propagated under repressing conditions (YMGlucas). Two-fold dilution series (from left to right) were prepared from cell suspensions (10^6 cells/ml) of four separate transformants, and replicated to YMGlucas (Glu) and YMGRcas (Gal) plates with a nail-block. Plates were incubated for 5 days at 30°C.

(C) Western Blot. Cells of the wild-type strain S150-2B bearing pRS316 (*CEN, URA3*; lane 1) or pAJY11 (*CEN, URA3, GAL1:PRP8*; lane 2) were grown under inducing conditions (YMGRcas) and splicing extracts prepared. Extracts (200µg protein per lane) were fractionated by SDS-PAGE through a 6% gel. Proteins were transferred to immobilon-P nylon membrane, and the Prp8p species visualised with anti-8.2 immune serum. Protein molecular weight standards (M) are indicated (size in kDa).

(D) Northern Blot. Total yeast RNA was extracted from S150-2B cells carrying pAJY11 (lanes 2 and 4) or pRS316 (lanes 1 and 3) grown in YMGlucas (repressing conditions; Glu) or YMGRcas (inducing conditions; Gal) liquid media. 60µg RNA (or 12µg for lane 4) was denatured and fractionated through a 1% agarose/formaldehyde gel, and blotted to hybond-N nylon membrane. The 5.8kb *Bam*HI fragment encoding most of *PRP8* was ^{32}P-labelled by random-primimg and hybridised to the RNA blot. The only hybridising band was in a position expected for *PRP8* mRNA (8kb), relative to 18S and 23S rRNA. The substantial mRNA accumulation with *GAL1:PRP8* is significant; however the blot was not hybridised to other probes to control for loading, and so a degree of variability (for example between lanes 1 and 2) cannot be regarded as significant.


A
Glu
Gal
23
34
23
34
pBM125
pBM-PRP8
B
Glu
Gal
1
2
3
4
5
6
7
8
C
S150-2B
M
pRS316
pAJY11
←Prp8p
205
116
97
1
2
D
Glu
Gal
vector
pAJY11
vector
pAJY11
1
2
3
4

colony purification, transformants were grown to stationary phase in YPGRA and streaked to YPGRA plates, and then individual colonies were replicated to YMGR plates with various growth supplements. Of 48 colonies tested, all retained the *URA3* marker (0/48 uracil auxotrophs) but only nine retained the *HIS3* marker (39/48 histidine auxotrophs). Thus pJDY13 could be lost, and seemed to be more unstable than pBM-PRP8 under the conditions of this experiment. As an experimental control the same procedure was carried out with the two vectors pBM125 and pRS316 (both *CEN, URA3*). With these, pJDY13 (*HIS3*) was stable but the *URA3* marker was frequently lost (14/48 uracil auxotrophs with pBM125 and 22/48 with pRS316). The strain generated from JDY8.05 by replacement of pJDY13 with pBM-PRP8 was named AJY8.08 (*ura3, his3, prp8::LEU2*, pBM-PRP8).

As shown above, *GAL1:PRP8* is repressed by glucose in a temperature-sensitive strain to levels at which *PRP8* expression is limiting for growth. From this it was expected that AJY8.08 would have a growth defect on glucose media. However several independent isolates of AJY8.08 grew on YMGlucas plates, and in liquid YMGlucas no growth defect could be detected compared to the isogenic control JDY8.02 (*ura3, his3, prp8::LEU2*, pY8500). Thus AJY8.08 in glucose has no splicing defect sufficient to significantly inhibit growth due to Prp8p deficiency. The actual extent of glucose repression with the *GAL1:PRP8* plasmids is apparently variable between strains.

4) Wild-type background: Expression of *PRP8* from pAJY11 was examined in a wild-type strain to determine induced mRNA and Prp8 protein levels. The wild-type strain S150-2B was used because this strain is known to yield strong expression of *GAL1* promoter-fusion constructs. Messenger RNA level was quantified by Northern blotting. In YMGRcas, pAJY11 (*GAL1:PRP8*) causes a massive accumulation of mRNA, roughly two orders of magnitude greater than that derived from the weakly-expressed chromosomal locus (Figure 3.2D, lanes 3 and 4). In this experiment, known quantities of yeast total RNA, calculated from A_{260}, were analysed. It later became clear that the mRNA population as a proportion of total yeast RNA is variable between preparations. In this experiment, blots were not hybridised to other probes to control for this, but variability of loading controls in other experiments is not greater than several-

fold. Similar *PRP8* mRNA accumulation was observed in independent yeast total RNA preparations (not shown), and thus the *PRP8* mRNA accumulation due to *GAL1:PRP8* detected here is regarded as significant.

The Western blot is presented in Figure 3.3C. In splicing extract prepared from S150-2B cells bearing pAJY11 (*CEN, URA3, GAL1:PRP8*), 3-5 fold more Prp8 protein was detected than in extracts from the same strain bearing pRS316 (*CEN, URA3*). This shows that Prp8 protein production due to induction of *GAL1:PRP8* is greater than that due to the chromosomal locus, confirming the genetic evidence.

In summary, pBM-PRP8 and pAJY11 are indeed functionally equivalent, and both enable the regulated overexpression of *PRP8*. pAJY11 was introduced into *prp8* null and temperature-sensitive backgrounds, and conferred the same phenotype as pBM-PRP8 (not shown). Transcription of *GAL1:PRP8* is induced by galactose, yielding far greater than wild-type levels of mRNA, and protein accumulation can be detected, although not in proportion to mRNA level. Repression is variable between strains but in general reduces expression to growth-limiting levels. The induced Prp8 protein level detected by Western blotting is not greater than levels caused by moderately raised gene-dosage (Jackson *et al.*, 1988), therefore the protein is probably not readily accumulated to high levels.

3.2.2 Construction of a Plasmid Constitutively Expressing *PRP8*.

As well as the regulated systems of *PRP8* expression described above, the plasmid pAJY12 (*CEN, URA3, PRP8*) was constructed. It is derived from a different vector to pY8500 (Table 3.1) and was used in the construction of truncation mutants. Plasmid pAJY12 encodes full-length, wild-type *PRP8* expressed from the promoter region of the chromosomal *PRP8* locus which was shown to express the gene at low levels (Jackson *et al.*, 1988). Its construction is presented in Figure 3.2B. JDY8.05 cells bearing pAJY12 are viable on glucose, demonstrating that pAJY12 can confer *PRP8* activity.

3.3 PLASMIDS BEARING DELETIONS OF THE *PRP8* N-TERMINAL DOMAIN.

This study analysed two truncated versions of Prp8 protein, one with 2 copies of the LP_nG motif (wild-type Prp8p has 4 copies), and one with none. The genes encoding these truncated proteins were expected to be recessive, and were named (with lower-case) *prp8ΔP2* (two copies of LP_nG) and *prp8ΔN* (no copies of LP_nG).

3.3.1 Construction of *prp8ΔN*.

This study used two plasmids which had been constructed previously, by J.D. Brown (unpublished). These were pJDY6 (2μ, *URA3, PRP8*) which contains full-length, wild-type *PRP8* under the trancriptional control of its own promoter, and pJDY7 (2μ, *URA3, prp8ΔN*) which was derived from pJDY6 by in-frame fusion of the *Nhe*I and *Xba*I sites. This removes the fragment encoding amino acids 4 to 79 inclusive from within the *PRP8* coding sequence, creating *prp8ΔN*. The protein encoded by *prp8ΔN* lacks the N-terminal domain (the four LP_nG motifs) but not the cluster of acidic amino acids at positions 81 to 84 (see Figure 3.1).

The same cloning procedure, an in-frame fusion of the *Nhe*I and *Xba*I sites, was carried out with pBM-PRP8, forming pAJY9 (*CEN, URA3, GAL1:prp8ΔN*). Thus pAJY9 and pJDY7 encode identical alleles (*prp8ΔN* in both cases), but under the transcriptional control of different promoters (see Table 3.1).

3.3.2 Construction of *prp8ΔP2*.

The strategy for partial deletion of the N-terminal domain, designed to create versions of *PRP8* containing 1, 2, or 3 copies of the LP_nG motif, was based on the polymerase chain reaction (PCR). A degenerate oligodeoxynucleotide primer was able to anneal to any one of the four strings of proline codons. This primer was used to amplify a nested set of fragments having 1, 2, 3 or 4 copies of LP_nG. Any of these

fragments could then be used to replace the N-terminus of wild-type *PRP8*. Figure 3.4 outlines this concept diagrammatically.

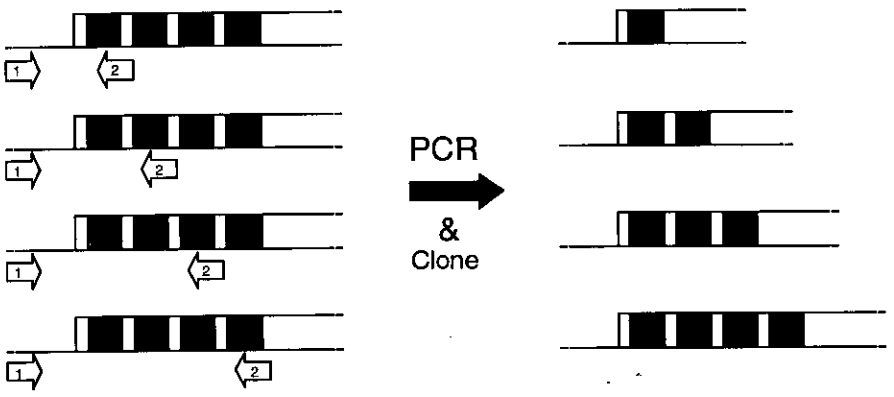


Figure 3.4: Concept for creation of a nested set of deletion mutations using a degenerate primer (2) able to anneal to any of the four LP_nG-encoding motifs (shaded boxes) in *PRP8* (open box).

Amplification was not attempted in a single step because of an envisaged problem: amplification would be biased towards shorter fragments of the nested set. Intrinsically, PCR is likely to amplify shorter targets more efficiently (if they have a better chance of complete synthesis), but this problem is compounded in synthesis of the nested set. This is because templates for the short fragments accumulate faster than those for longer fragments (ie the $(LP_nG)_4$ product is a template for $(LP_nG)_1$, but the $(LP_nG)_1$ product is not a template for $(LP_nG)_4$). To circumvent this, synthesis of the nested set and amplification were separated into two steps. The details of this approach are presented in Figure 3.6. In the first step, the polymerase reaction contained only one primer (A969) annealed to any one of the four strings of proline codons at the N-terminus. A969 incorporated degeneracy for two or three bases at four positions, and thereby was able to form regions of base-pairing at least 12 bases long with any repeat (Figure 3.5). A969 was annealed to a molar excess of the template (pBM-PRP8), to prevent multiple primers annealing to individual molecules of the template, and extended with Taq polymerase to create the nested set (as single-stranded DNA). In the second step these products were amplified with a pair of primers. Selectivity was achieved by using a truncated version of A969, so that only the nested set and not the *PRP8* plasmid was amplified. This primer also contained a 9 nucleotide stretch of 100% G or C nucleotides at its 5' end (a GC-clamp) to improve PCR yield. The second

primer was complementary to the *GAL1* promoter, such that the product contained 600 bp of the *GAL1* promoter region. Even with these refinements, the $(LP_nG)_1$ and $(LP_nG)_2$ fragments were more readily detectable than $(LP_nG)_3$ and $(LP_nG)_4$. This bias may be due to the considerations above; alternatively the bias may derive from oligonucleotide synthesis, if at degenerate positions different bases were not added with equal efficiency. Consistent with the latter possibility, less $(LP_nG)_1$ than $(LP_nG)_2$ fragment was obtained (not shown).

To improve yields prior to cloning, fragments were recovered from gels and reamplified under the same conditions as above, before being treated with Klenow fragment to create blunt ends, and cloned into the *Eco*RV site of the sequencing vector pBluescript KS^+. A cloned $(LP_nG)_2$ fragment (pAJ3) was sequenced and used to replace the N-terminal domain and upstream region of pY8500 (*CEN, URA3, PRP8*). This created pAJY1 (*CEN, URA3, GAL1:prp8ΔP2*) with two instead of the usual four N-terminal repeats (see Figure 3.7). This truncated allele of *PRP8* (*prp8ΔP2*) encodes a protein lacking amino acids 29 to 79 inclusive. To place this allele under the transcriptional control of the *PRP8* promoter, a fragment was inserted into pAJY12 (*CEN, URA3, PRP8*) to create pAJY13 (*CEN, URA3, prp8ΔP2*). Thus pAJY1 and pAJY13 both encode the same allele (*prp8ΔP2*) but with different promoters; they are described in Table 3.1.

Figure 3.5 Complementarity between the degenerate oligodeoxynucleotide A969 and each string of proline codons in *PRP8*.

Figure 3.6: Strategy for Partial Deletion of the N-terminal Domain of *PRP8*.

Step 1: A limiting quantity of the degenerate primer A969 was annealed to pBM-PRP8 and extended with thermalase (IBI). A969 is able to anneal to any of the four N-terminal repeats; for simplicity it is shown in this figure annealed to the second repeat. The reaction volume (50µl) contained 5µg pBM-PRP8, 1.5ng A969 and 1U thermalase. After denaturation (93°C, 10min) it was subjected to 5 cycles of 93°C (1min), 43°C (1min), 70°C (1min).

Step 2: The products of primer extension (step 1) were amplified in a standard PCR reaction with primers 989S and 229G, 1U thermalase, and 1µl of the step 1 reaction as template. Four products of the expected sizes were obtained, one of which was blunt-end cloned to the *Eco*RV site of pBluescript KS^+ (creating pAJ1) and identified as a fragment with two copies of the N-terminal repeat by sequencing. The cloning steps by which a gene encoding Prp8p with 2 copies of the LP_nG motif (*prp8ΔP2*) was constructed are presented in Figure 3.7.

Key for Restriction Enzymes: N,*Nhe*I; X,*Xba*I; O,*Xho*I.

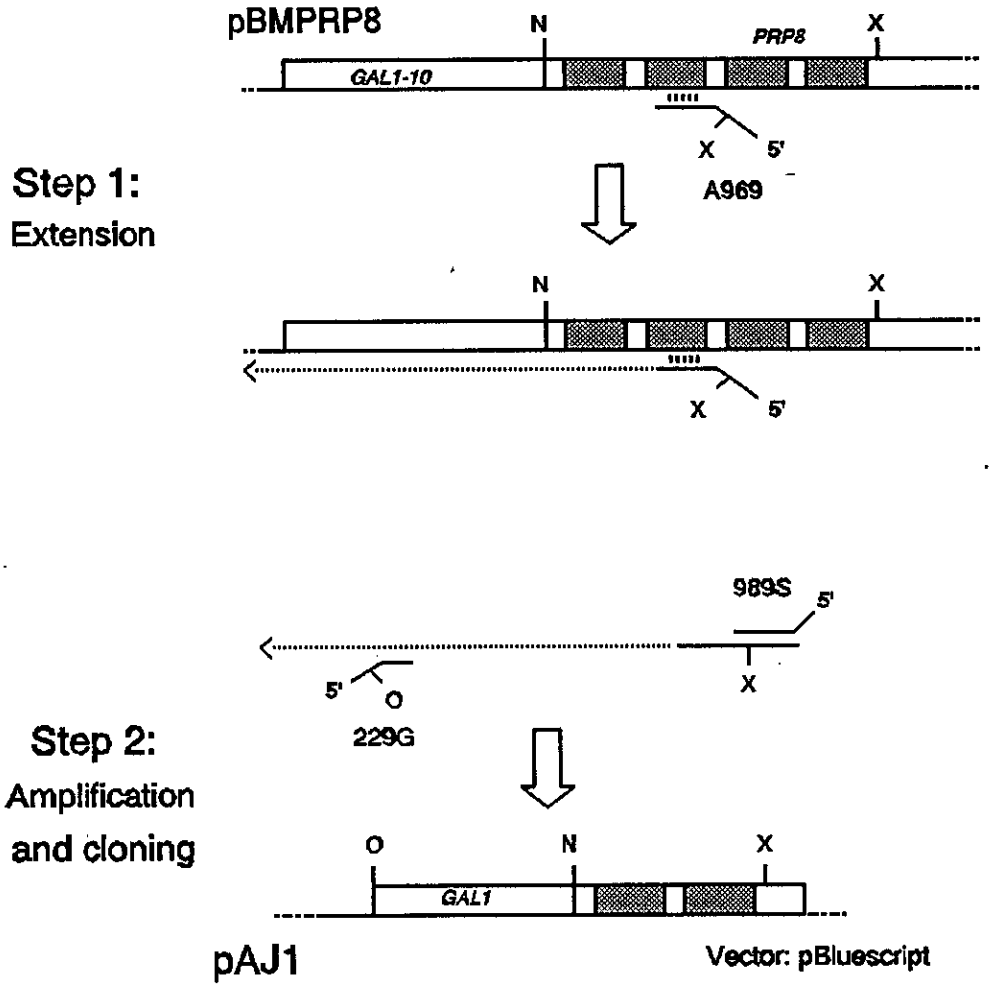

pBMPRP8
N
PRP8
X
GAL1-10
X
5'
A969
Step 1:
Extension
N
X
X
5'
989S
5'
5'
O
X
229G
Step 2:
Amplification
and cloning
O
N
X
GAL1
pAJ1
Vector: pBluescript

Figure 3.7: Construction of *prp8ΔP2*.

Plasmid pAJ1 was described in Figure 3.6; it contains a fragment amplified by PCR (box) consisting of the N-terminal end of *PRP8* (two copies of LP_nG; shaded box) and the *GAL1* promoter. Reconstruction of this fragment with the remainder of *PRP8* was carried out as follows:

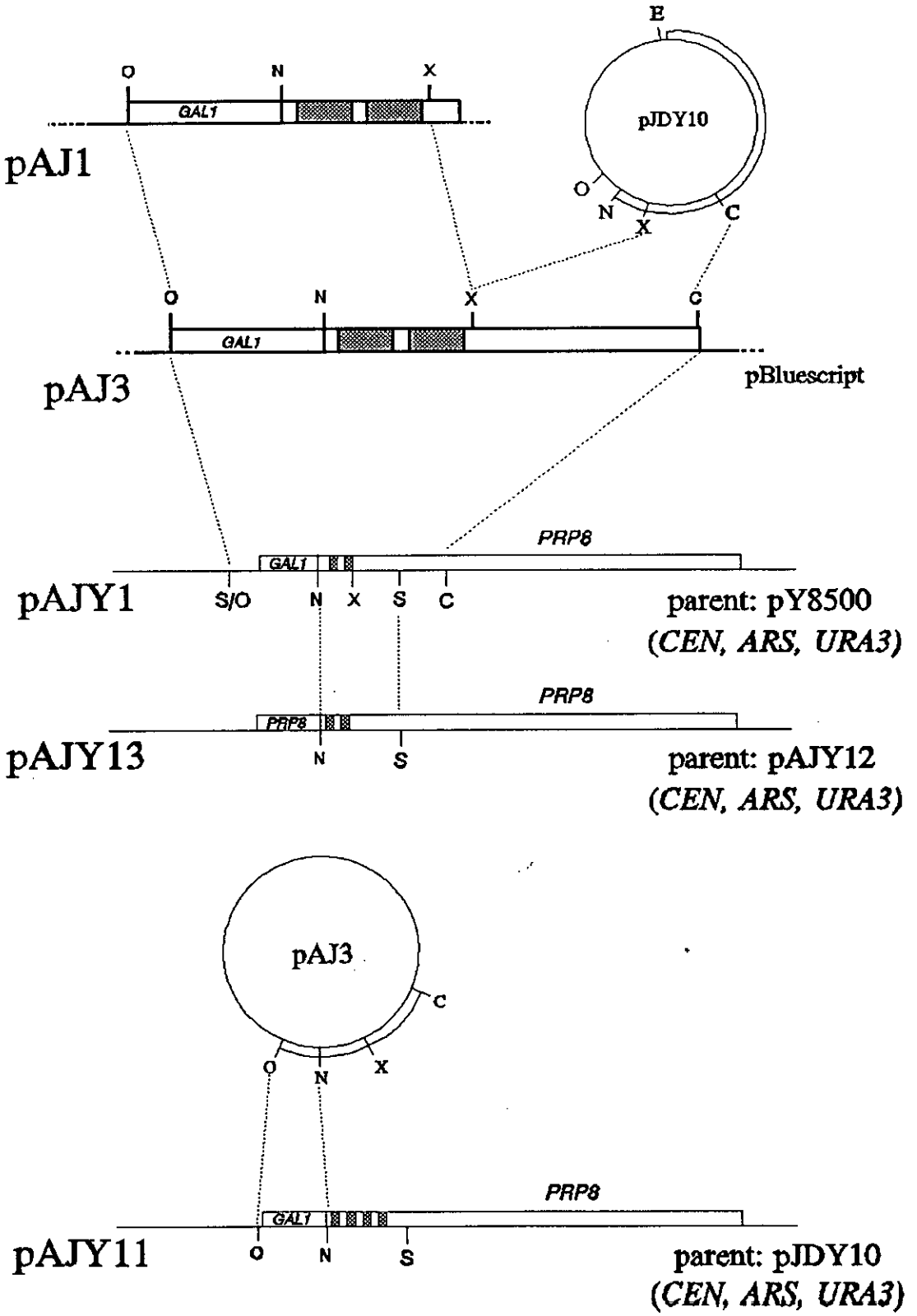
The 2.47kb *Xba*I-*Sac*I *PRP8* fragment (from pJDY10) was inserted into the pBluescribe KS^- polylinker. Next, the *Xho*I-*Xba*I fragment from pAJ1 was inserted, to create pAJ3.

pAJY1 (*CEN, URA3, GAL1:prp8ΔP2*) was created by replacing the *Sal*I-*Sac*I region of pY8500 (*CEN, URA3, PRP8*) with the *Xho*I-*Sac*I fragment (containing the *GAL1* promoter, two copies of the LP_nG motif, and 2.5kb of *PRP8*) from pAJ3.

pAJY13 (*CEN, URA3, prp8ΔP2*) was created by inserting the *Nhe*I-*Sal*I fragment of pAJY1 into pAJY12 (*CEN, URA3, PRP8*).

pAJY11 (*CEN, URA3, GAL1:PRP8*) was created by replacing the *XhoI-NheI* fragment of pJDY10 (*CEN, URA3, GAL1UAS:PRP8*) with the *XhoI-NheI* fragment of pAJ3.

Key for Restriction Enzymes: E, *Eag*I; S, *Sal*I; C, *Sac*I; and as Figure 3.6. S/O: fusion of *Sal*I and *Xho*I half-sites.


pAJ1
O
N
X
GAL1
pJDY10
E
O
N
X
C
pAJ3
O
N
X
C
GAL1
pBluescript
pAJY1
PRP8
GAL1
S/O
N
X
S
C
parent: pY8500
(CEN, ARS, URA3)
pAJY13
PRP8
PRP8
N
S
parent: pAJY12
(CEN, ARS, URA3)
pAJ3
C
O
N
X
pAJY11
PRP8
GAL1
O
N
S
parent: pJDY10
(CEN, ARS, URA3)

3.4 Discussion.

This chapter describes the construction of *PRP8* overexpression systems, and of truncation mutants of *PRP8*. These were preliminary steps to investigating the function in the splicing process of the proline-rich N-terminal domain, which was known to be essential for viability under normal conditions. The two new *GAL1:PRP8* constructs, pBM-PRP8 and pAJY11, behaved similarly in complementing *prp8* temperature-sensitive and null strains. The *GAL1* region of pAJY11 was generated by PCR and has not been fully sequenced, but considering pAJY11 mimics pBM-PRP8 in which the promoter sequence is known, any mutation(s) due to PCR error cannot significantly affect *GAL1* induction or repression. The production of *PRP8* mRNA from pAJY11 under inducing and repressing conditions is consistent with this. These systems are designed to facilitate further directed mutagenesis of *PRP8*. In particular pAJY11 contains unique *Nhe*I, *Eag*I and *Xho*I sites, enabling *PRP8* or *GAL1:PRP8* to be transferred as a cassette to vectors with alternative markers, such as the pRS series (Sikorski and Hieter, 1989).

The lethality of wild-type *PRP8* overexpression in the *spp81(ded1)* background is of interest because the nature of the relationship between *PRP8* and *DED1* is still poorly understood. The genetic interaction between the two is well established (Jamieson *et al.*, 1991), but its functional significance is unknown, as *DED1* does not seem to have a role in splicing (S. Maccallum and J.D. Beggs, unpublished data). *DED1* has been isolated as a suppressor of defects in other cellular functions (both as a mutant and a high-copy suppressor). Recent experimental evidence has demonstrated that *ded1* mutants are defective in translational initiation at the restrictive temperature (T.-H. Chang, R.-Y. Chuang, and P. Weaver, unpublished data) and this is possibly the basis of suppression. The ability to conditionally overexpress *PRP8* in *spp81* cells offers the opportunity to study further the interaction between *PRP8* and *DED1*, and to determine if the toxicity is associated with a defect in pre-mRNA splicing and/or translational initiation.

The *GAL1:PRP8* system produces large quantities of mRNA, but does not accumulate high levels of protein, either due to inefficient translation or protein instability. Jackson (1987) calculated that PRP8 has almost no preference for those

codons which predominate in highly-expressed yeast genes, as expected for a yeast gene expressed at low levels. The corrections which have been made to the compiled sequence since that time do not significantly affect this conclusion. Therefore, availability of rare tRNA species may limit the translation of *PRP8* mRNA, and protein production from the *GAL1:PRP8* system. Polypeptide components of other ribonucleoprotein particles such as ribosomal proteins are characteristically unstable except when assembled into ribonucleoprotein particles (Wittekind *et al.*, 1990; Moritz *et al.*, 1990; Woolford and Warner, 1991). If this is true also of Prp8p, co-overexpression of *PRP8* with U5 snRNA and/or other components of the U5 snRNP may be required for higher-level accumulation of Prp8p. The *GAL1:PRP8* plasmids in the null background of JDY8.05 supported apparently normal growth on glucose (repressing conditions). This suggests that glucose repression is unusually weak in the JDY8.05 strain. A degree of repression must occur, because this strain was used to deplete Prp8p (Brown and Beggs, 1992). Presumably, depletion of Prp8p by glucose repression of JDY8.05 was only possible because Prp8p levels were low even in galactose (inducing conditions).

The PCR strategy used to delete half of the *PRP8* N-terminal domain is generally applicable as a rapid and cost-effective means to create numerous, progressive deletions of long repetitive domains. As the proline codon is G/C-rich (proline=CCx), repetitive proline-rich domains are especially suited. During the course of creating the partial deletions of the N-terminal domain, it was discovered that *prp8*ΔN which lacks the entire N-terminal domain was functional under conditions of overexpression (this is described in the next chapter). For this reason the other two *PRP8* truncations (*prp8ΔP1* and *prp8ΔP3* in this nomenclature) which this strategy was designed to yield were not pursued.

Chapter 4.

The *In vivo* Phenotypes of *PRP8* Truncation Mutants.

4.1 Introduction.

The preceding chapter described the construction of deletion mutants *prp8ΔP2* and *prp8ΔN*, with two or no copies of the LP_nG motif respectively. An objective of this work was to generate a mutant phenotype which could indicate the function of the N-terminal domain. Considering that deletion of the N-terminal domain knocks out function, two possible phenotypes were envisaged. Firstly, the loss-of-function *prp8ΔN* might interfere with wild-type function when overexpressed (a dominant negative phenotype) especially if it retained the ability to assemble into spliceosomal complexes but was blocked at some later stage of the pathway at which the N-terminus was functional. Dominant negative alleles of another splicing factor gene, the DEAD-box putative RNA helicase *PRP2* have been highly revealing (Plumpton *et al.*, 1994). Secondly, partial deletion of the N-terminal domain might result in conditional function only at low (or high) temperatures. Having obtained cold- or heat-sensitive alleles, strategies to identify factors interacting with the N-terminal domain could be designed. This chapter describes the introduction of *prp8ΔN* and *prp8ΔP2* into yeast, and the phenotypes which result.

4.2 Dominant Phenotypes.

To determine if overexpression of the truncation mutants was inhibitory to growth, rates of growth of S150-2B (wild-type) cells bearing pAJY1 (*CEN, URA3, GAL1:prp8ΔP2*) and pAJY9 (*CEN, URA3, GAL1:prp8ΔN*) were analysed on agar plates or in liquid media. Under both inducing (YMGRcas; galactose) and repressing (YMGlucas; glucose) conditions, no difference could be detected compared to transformants containing vector (pBM125 (*CEN, URA3*); data not presented). This indicated that overexpression of the truncation mutants does not interfere with wild-type function; the truncation mutants are not dominant negative.

4.3 Complementation of the *prp8-1* Allele.

The complementation of temperature-sensitive *prp8* strains, which are defective for *PRP8* activity at the restrictive temperature, is a primary test of function of novel alleles created by directed mutagenesis techniques. Temperature-sensitive SPJ8.31 (*prp8-1*) cells bearing a series of plasmid constructs expressing *prp8ΔN* or *prp8ΔP2* were tested for growth at the restrictive temperature (36°C). Both truncation mutants complemented *prp8-1* temperature sensitivity (Figure 4.1A). This was unexpected because it had been shown previously (and is confirmed later in this chapter) that *prp8ΔN* does not support growth of a *prp8* null allele (J.D. Brown, unpublished). This was shown using the construct pJDY7 (2μ, *URA3, prp8ΔN*). This suggested that the requirement for complementation of *prp8* temperature-sensitive and null mutants was fundamentally different.

One possible explanation of this was that *prp8ΔN* borne on a plasmid rescued the *prp8-1* point mutation in the chromosomal allele by gene conversion. This is theoretically possible because plasmids such as pJDY7 contain wild-type *PRP8* sequence spanning the position of *prp8-1* (Hodges *et al.*, 1995). However, *prp8-1* rescue by gene conversion has been demonstrated to occur at low frequency in this strain (Jackson, 1987), and multiple transformants tested in this experiment were all viable at 36°C. Also, cells bearing pAJY9 (*CEN, URA3, GAL1:prp8ΔN*) were only viable at 36°C on galactose; gene-conversion to both rescue the *prp8-1* allele and place it under *GAL1* control is highly unlikely. Therefore, this result suggests that *prp8ΔN* confers a degree of *PRP8* function in temperature-sensitive strains, contrary to its behaviour in the null background.

A second formal possibility was that the truncated protein protects the product of the *prp8-1* allele from heat inactivation. This was investigated by determining which version of Prp8p was present in snRNP complexes, by immunoprecipitation. SPJ8.31 (*prp8-1*) cells bearing pJDY7 (2μ, *URA3, prp8ΔN*) were grown in YMGlucas to select for plasmid maintenance and at the restrictive temperature (36°C), and a splicing extract was prepared. The extract was active in standard splicing assays (data not

Figure 4.1: Function of Truncation Mutants in a Temperature-sensitive *prp8-1* Strain.

(A) SPJ 8.31 (*prp8-1*) is complemented by the *prp8ΔN* truncation mutant.

SPJ8.31 cells bearing various *URA3* plasmids were suspended in sterile distilled water to a density of approx. 10^6 cells/ml, and transferred in 20µl spots to plasmid-selective plates for growth under inducing (YMGRcas) or repressing conditions (YMGlucas). Weak complementation by pBM-PRP8 under repressing conditions is believed to be due to incomplete repression. The *URA3* plasmids were:

(1) vector:	pBM125	(5) *PRP8*:	pJDY6
(2) *GAL1:PRP8*:	pBM-PRP8	(6) *prp8ΔN*:	pJDY7
(3) *GAL1:prp8ΔP2*:	pAJY1	(7) *PRP8*:	pAJY12
(4) *GAL1:prp8ΔN*:	pAJY9	(8) *prp8ΔP2*:	pAJY13

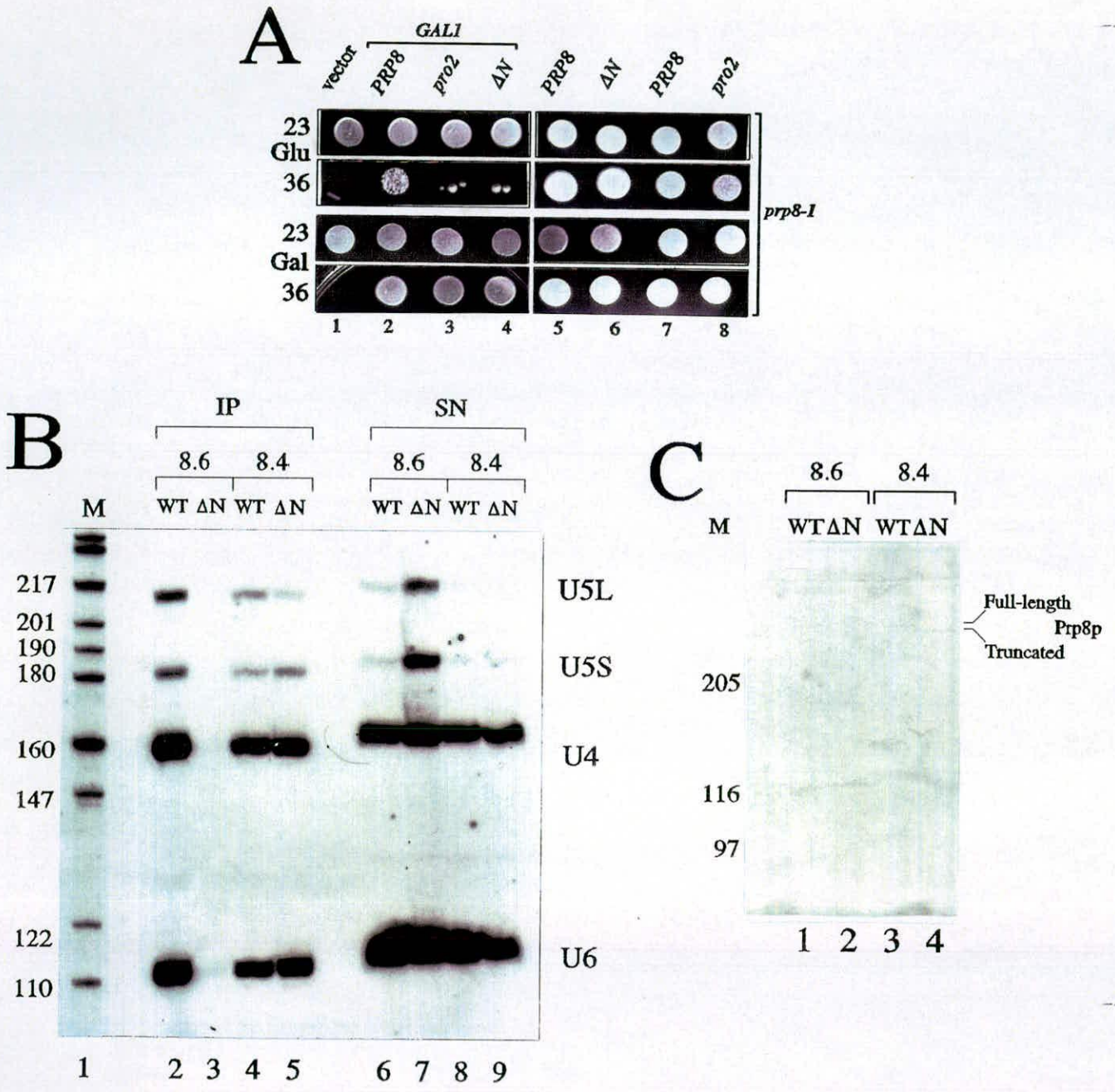
(B): U5 snRNP complexes in SPJ8.31 (pJDY7) extract lack the 8.6 epitope.

Cultures of SPJ8.31 (*prp8-1*) cells bearing pJDY7 (2µ, *URA3*, *prp8ΔN*) were grown at the restrictive temperature for *prp8-1* (36°C). The growth medium was YMGlucas (lacking uracil) to select for plasmid maintenance. At mid-log phase, splicing extracts were prepared. Mock splicing reactions were assembled with S150-2B (wild-type) extract or the SPJ8.31 (pJDY7) extract (1.4mg protein per mock reaction), and subjected to immunoprecipitation with anti-8.4 and anti-8.6 antibodies. To analyse both precipitated proteins and snRNAs, the procedure was scaled up two-fold. 50% of each precipitate was analysed for snRNA content (lanes 1 to 4) by deproteinisation, followed by electrophoresis through a 6% denaturing polyacrylamide/urea gel. SnRNAs were electroblotted to hybond-N nylon membrane and hybridised to end-labelled deoxynucleotides complementary to U4, U5 and U6 snRNA. 10% fractions of the supernate remaining after precipitation (lanes 5 to 8) were treated likewise.

The two U5 species, U5L and U5S are both transcribed from the *snR7* gene and differ in length at the 3' end. Both are thought to be functional.

(C): Prp8p produced from the truncation mutant *prp8ΔN* enters spliceosomal complexes in strain SPJ8.31 (*prp8-1*).

Splicing complexes were immunoprecipitated from SPJ8.31 (pJDY7) and wild-type extracts with anti-8.6 and anti-8.4 antibodies (see (B) above). 50% of each precipitate was boiled with SDS protein loading buffer, and extracted proteins fractionated by electrophoresis through a 6% SDS-polyacrylamide gel. Proteins were electroblotted to Immobilon-P membrane and decorated with anti-8.4 serum. Bands are faint; the putative truncated and full-length Prp8p bands are indicated.


A
GAL1
vector
PRP8
prp2
ΔN
PRP8
ΔN
PRP8
prp2
23 Glu
36
23 Gal
36
prp8-1
1 2 3 4 5 6 7 8
B
IP
SN
8.6 8.4
8.6 8.4
M
WT ΔN WT ΔN
WT ΔN WT ΔN
217
201
190
180
160
147
122
110
U5L
U5S
U4
U6
1 2 3 4 5 6 7 8 9
C
8.6 8.4
M
WTΔN WTΔN
205
116
97
Full-length Prp8p
Truncated
1 2 3 4

4.4 Complementation of the Null Background.

As with temperature-sensitive strains, the complementation of a *prp8* null allele is a test of function of novel *PRP8* alleles. The *prp8*-null genetic background also enables the isolation of strains in which the novel allele is the sole gene copy, by plasmid shuffle. The results of introducing the truncation mutants into the strain JDY8.05 which contains the *prp8* null allele are summarised in Table 4.1.

4.4.1 *prp8ΔP2*

Strain JDY8.05 (*ura3, his3, prp8::LEU2*, pJDY13) was transformed with pAJY13 (*CEN, URA3, prp8ΔP2*) or pAJY12 (same except *PRP8*) and transformants selected on YMGR lacking both uracil and histidine. Transformants therefore contained two plasmids: the *URA3* plasmid and pJDY13 (*CEN. HIS3, GAL1:PRP8*). The double transformants were viable on glucose (YMGlu plates) unlike the vector (pRS316) control, demonstrating that the plasmid which produces Prp8 protein with two copies of LP_nG confers a degree of *PRP8* function (Figure 4.2A).

The double transformant bearing pAJY13 (*prp8ΔP2*) and pJDY13 (*GAL1:PRP8*) was viable on glucose only at 23°C. The control (pAJY12 (*PRP8*)/pJDY13 (*GAL1:PRP8*) double transformant) grew at both 23°C and 34°C, as expected (Figure 4.2A). Plasmid shuffle was carried out by relaxing selection for pJDY13 (*HIS3*) while maintaining selection for the *URA3* plasmid (on glucose, in this case), and then screening for a histidine auxotroph. From the control pAJY12 (*URA3, PRP8*)/pJDY13 (*GAL1:PRP8*) double transformant, pJDY13 was cured readily. In contrast histidine auxotrophs could not be isolated on histidine-containing medium from the pAJY13 (*URA3, prp8ΔP2*)/pJDY13 (*GAL1:PRP8*) double transformant, suggesting that there is a selection against loss of pJDY13. It may be that cure of pJDY13 is possible, but occurs at very low rates. There is no selection for the His^- phenotype (unlike the Ura^- phenotype which can be selected with 5-fluoro-orotic acid) so no selection for pJDY13 cure in this plasmid shuffle can be set up. The temperature sensitivity and resistance to pJDY13 cure of the double transformant suggest that removal of two of the four LP_nG repeats from Prp8p reduces activity of the protein.

Table 4.1: A Summary of Results.

Promoter ⇨ / Gene ↓ / Test ⇨	*PRP8* Upstream Region			*GAL1* promoter		
	Viability of JDY8.05+plasmid on glucose			Growth on galactose after plasmid shuffle		
	Plasmid			Plasmid		
PRP8 (wild-type)	pAJY12 pJDY6	✓	Cured of pJDY13	pBMPRP8 pAJY11	✓	Not Glucose-repressible
prp8ΔP2 (partial truncation)	pAJY13	✓ (ts)	Not Cured of pJDY13	pAJY1	✓	Glucose-repressible
prp8ΔN (full truncation)	pJDY7	X	Not Cured of pJDY13	pAJY9	✓ (ts)	Glucose-repressible AJY8.07
Vector control	pRS316 YEp24	X	Not Cured of pJDY13	pBM125	ND	Not Cured of pJDY13

ts: temperature-sensitive at 34°C X :Not Viable ✓ :Viable ND :Not determined

Figure 4.2: *PRP8* Truncation Mutants in Null Backgrounds.

(A): JDY8.05 (pAJY13) cells are temperature-sensitive.

JDY8.05 (*prp8::LEU2*, pJDY13) cells bearing *URA3* plasmids were spotted to plates (supplemented with adenine and tryptophan) lacking both uracil and histidine. This maintained selection for both the *URA3* plasmid and pJDY13 (*HIS3*). Cells were incubated for 5 days under either inducing (GAL; YMGR plates) or repressing (Glu; YMGlu plates) conditions at 23°C and 34°C. The *URA3* plasmids were pRS316 (*CEN, URA3*), pAJY13 (*CEN, URA3, prp8ΔP2*) and pAJY12 (*CEN, URA3, PRP8*).

(B): JDY8.05 (pAJY1) cells are galactose-dependent.

pJDY13 was cured from JDY8.05 cells bearing either pAJY1 (*GAL1:prp8ΔP2*) or pAJY11 (*GAL1:PRP8*) by plasmid shuffle, carried out as described previously (Section 3.2.2.3) except transformants were grown under conditions (YMGR plus histidine, adenine and tryptophan) which maintained selection for the *URA3* plasmid. Histidine auxotrophs (12 isolates bearing pAJY11, top; 26 isolates bearing pAJY1, bottom) were transferred as suspensions to YMGlu (Glu) or YMGR (Gal) plates (supplemented with histidine, adenine and tryptophan) and incubated at 23°C or 34°C.

(C): AJY8.07 is associated with a splicing defect at restrictive temperature.

AJY8.07 is derived from JDY8.05 (pAJY9) by cure of pJDY13, and was temperature-sensitive at 34°C. AJY8.07 or JDY8.05 cells were grown continuously in YPGRA at 23°C (lanes 1 and 3) or heat-shocked from 23°C to 30°C (lane 4) or 34°C (lanes 2 and 5) or transferred (without changing temperature) from YPGRA to YPDA (Glu) to cause depletion of Prp8p. Total yeast RNA was prepared 6 hours after the time of heat-shock/media transfer. 60μg total RNA was fractionated through 1.4% agarose/formaldehyde by electrophoresis, blotted to hybond-N nylon membrane, and hybridised to a ^{32}P-labelled DNA fragment encoding the intron-containing *RP28* gene. Messenger RNA (mRNA) and pre-mRNA are indicated.

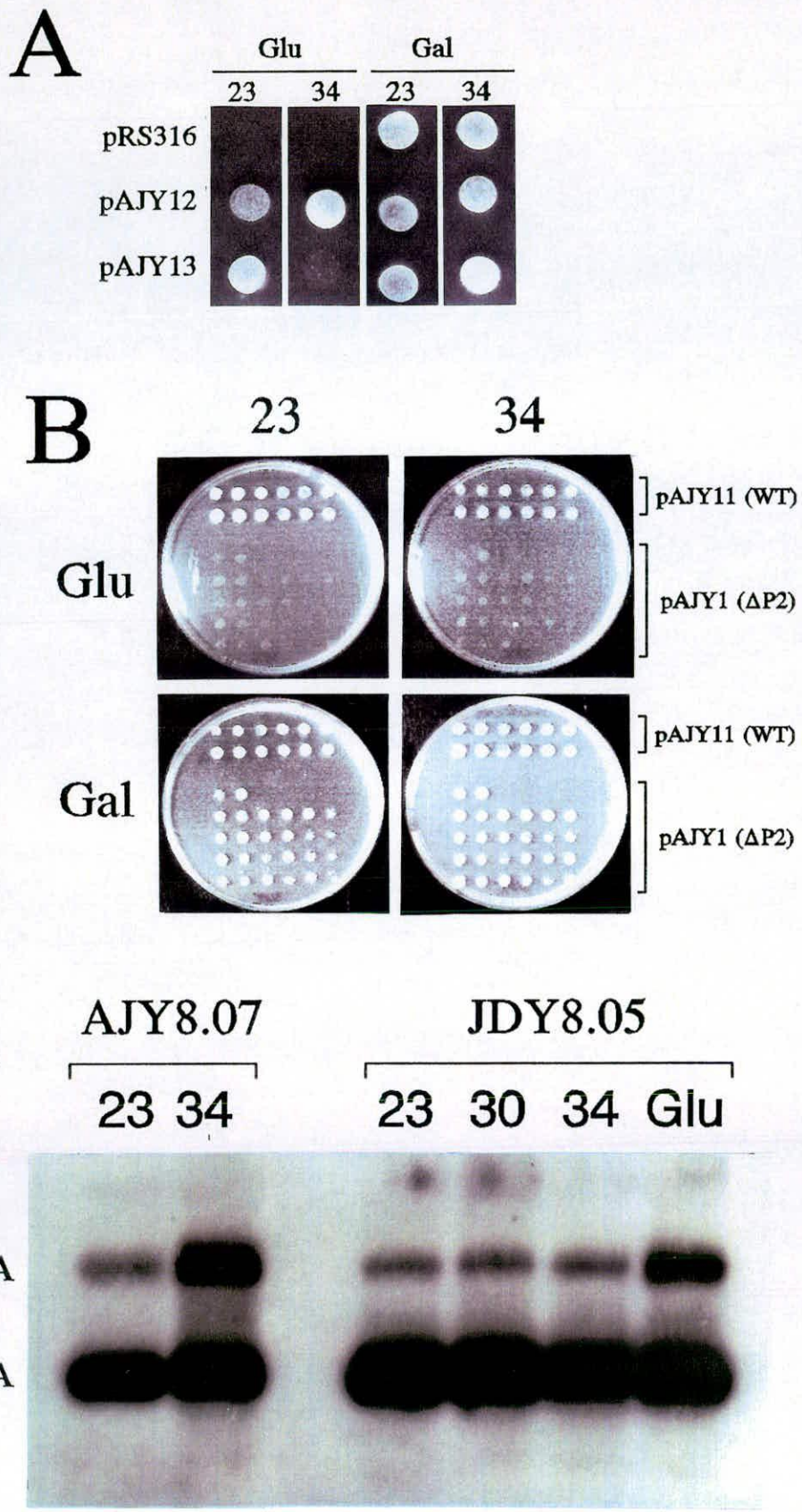
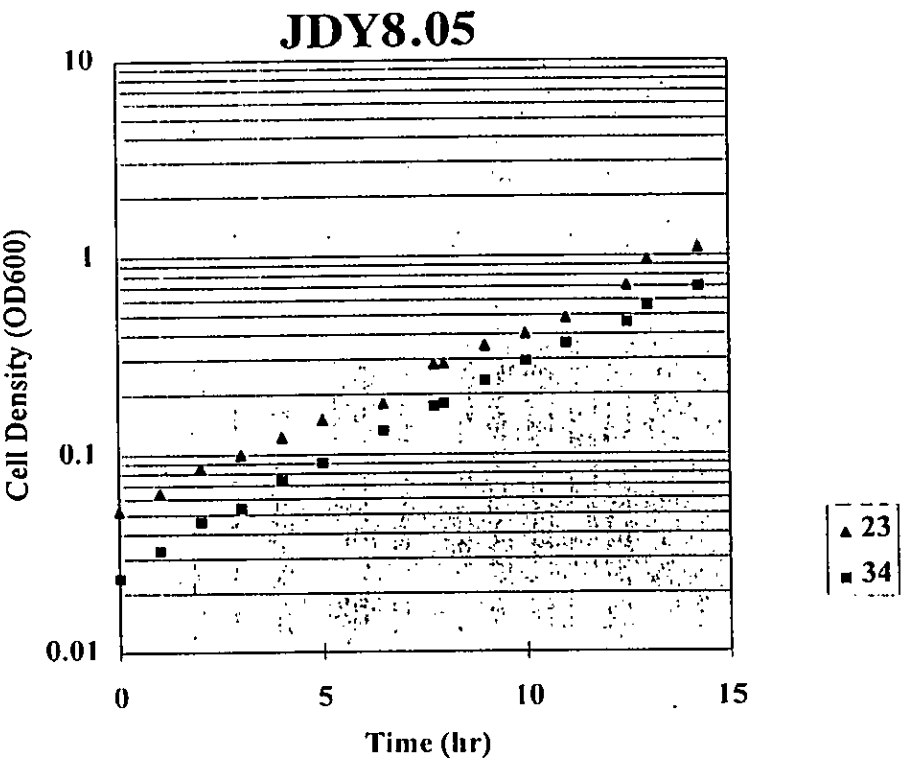

A
Glu
Gal
23
34
23
34
pRS316
pAJY12
pAJY13
B
23
34
Glu
Gal
pAJY11 (WT)
pAJY1 (ΔP2)
pAJY11 (WT)
pAJY1 (ΔP2)
C
AJY8.07
JDY8.05
23 34
23 30 34 Glu
pre-mRNA
mRNA

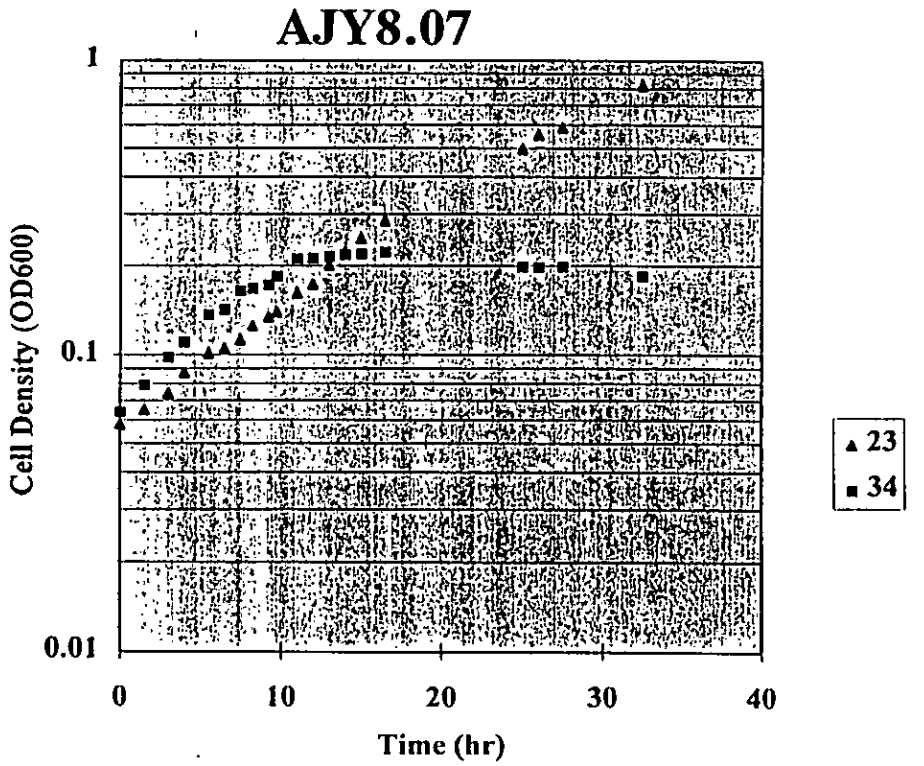
Figure 4.3: Growth Curves of JDY8.05 and AJY8.07 in YPGRA.

Mid-log cultures of strains JDY8.05 and AJY8.07 (23°C) were mixed (at t=0) with equal volumes of pre-warmed, pre-aerated medium at 23°C or 45°C and incubation continued at 23°C or 34°C respectively. OD_{600} was measured at intervals, and where necessary cultures were diluted with appropriate medium to maintain conditions appropriate for logarithmic growth. Aliquots of cultures were harvested after 6hr to prepare RNA for Northern blot in Figure 4.2C. Scales of both axes of the two graphs differ.

presented), and must therefore contain active U5 snRNP particles. Mock splicing reactions were incubated with PAS-coupled anti-Prp8p antibodies, directed against either the N-terminal domain (anti-8.6 antibodies) or against an internal region (anti-8.4 antibodies). Anti-8.4 antibodies coprecipitated the three snRNAs of the tri-snRNP particle, U4, U5, and U6 (Figure 4.1B, lane 5), confirming that in the SPJ8.31 (pJDY7) extract Prp8p was a component of functional U5 snRNP particles. Anti-8.6 antibodies were able to precipitate U5 snRNP particles from a wild-type control extract (lane 2), but not from the SPJ8.31 (pJDY7) extract (lane 3), although the supernate remaining after precipitation contained full-length U5 snRNA (lane 7). Therefore snRNP particles in this extract lack the 8.6 epitope. The same immunoprecipitates were also examined by Western blotting. The blot was decorated with anti-8.4 serum to visualise both truncated and full-length Prp8p. Quantities of Prp8p recovered were low, and close to the limits of detectability. Both anti-8.4 and anti-8.6 precipitates from the wild-type control extract contained co-migrating faint bands at the expected position for wild-type Prp8p (Figure 4.1C, lanes 1 and 3). The anti-8.4 precipitate from SPJ8.31 (pJDY7) contained a faster-migrating faint band at a position expected for truncated Prp8p (lane 4); the anti-8.6 precipitate from SPJ8.31 (pJDY7) contained no detectable Prp8p band (lane 2). Therefore the only detectable Prp8p species in U5 snRNP particles in this SPJ8.31 (pJDY7) extract was the truncated product of *prp8ΔN*. No full-length Prp8p was detected in the Western blot of the SPJ8.31 (pJDY7) anti-8.4 precipitate (lane 4), but it is possible that the protein product of *prp8-1* is precipitated at lower levels. However, the Northern demonstrates that it is not in snRNP particles. Therefore, protection of *prp8-1* against heat-inactivation is unlikely to be the basis of SPJ8.31 transformant viability, and under these conditions the truncated protein is apparently functional.

In conclusion, plasmid pJDY7 (*prp8ΔN*) does not support growth of the null mutant but does support growth of a temperature-sensitive mutant. Thus the requirements for complementation of the two types of *prp8* mutant must be fundamentally different.


JDY8.05
10
1
0.1
0.01
Cell Density (OD600)
0
5
10
15
Time (hr)
23
34


AJY8.07
1
0.1
0.01
Cell Density (OD600)
0
10
20
30
40
Time (hr)
23
34

Plasmid shuffle from a pAJY1 (*CEN, URA3, GAL1:prp8ΔP2*)/pJDY13 (*CEN, HIS3, GAL1:PRP8*) double transformant, in the same background as above (JDY8.05) resulted in histidine auxotrophs cured of pJDY13 (Figure 4.2B). This indicates that in a strain where *prp8ΔP2* is overexpressed, the truncation mutant with two copies of LP_nG can carry out all the activities of full length wild-type Prp8p. Also, overexpression of the mutant allele seems to overcome the effect of truncation. Figure 4.2B shows that the strain with the chromosomal *prp8* null allele and bearing pAJY1 (*GAL1:prp8ΔP2*) is not temperature-sensitive, but is only viable on galactose. The equivalent strain bearing pAJY11 (*GAL1:PRP8*) is not galactose dependent. This is consistent with the results above, and suggests that the *prp8ΔP2* truncation mutant exhibits an intermediate level of activity between the full-length, wild-type gene, and the *prp8ΔN* full truncation mutant.

4.4.2 *prp8ΔN*

JDY8.05 (*ura3, his3, prp8::LEU2*, pJDY13) cells bearing pJDY7 (2μ, *URA3, prp8ΔN*) were not viable on glucose, and attempts at plasmid shuffle did not yield histidine auxotrophs. This was a repeat of the experiment of J.D. Brown, and confirms his finding that *prp8ΔN* under the transcriptional control of the *PRP8* promoter does not complement the null strain.

Cure of pJDY13 was also attempted with JDY8.05 cells bearing pAJY9 (*CEN, URA3, GAL1:PRP8ΔN*). After the above result pJDY13 was expected to be stable in this strain, but histidine auxotrophs were isolated. They were designated AJY8.07 (putative genotype: *his3, ura3, prp8::LEU2*, pAJY9). This strain was galactose-dependent, grew appreciably slower than JDY8.05 at 23°C, and was temperature-sensitive on agar plates at 34°C.

If pJDY13 had been lost from AJY8.07 it was predicted that truncated Prp8p would be present in spliceosomal complexes. This was demonstrated by immunoprecipitation. Anti-8.4 serum precipitated U4, U5, and U6 snRNA from mock reactions with AJY8.07 extract; anti-8.6 serum did not (Figure 4.4B). In Western blots of the same extract, a species was detected which migrates faster than Prp8p, and which

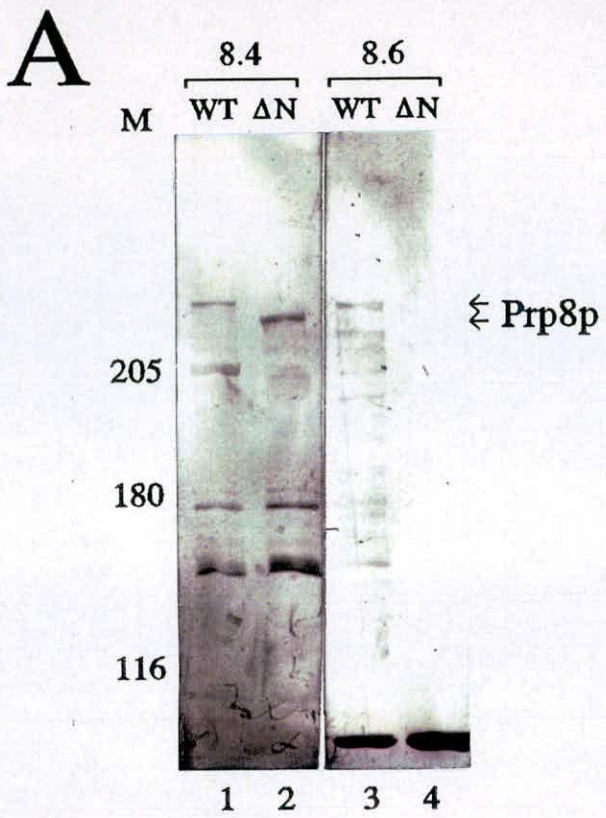
Figure 4.4: Analysis of AJY8.07.

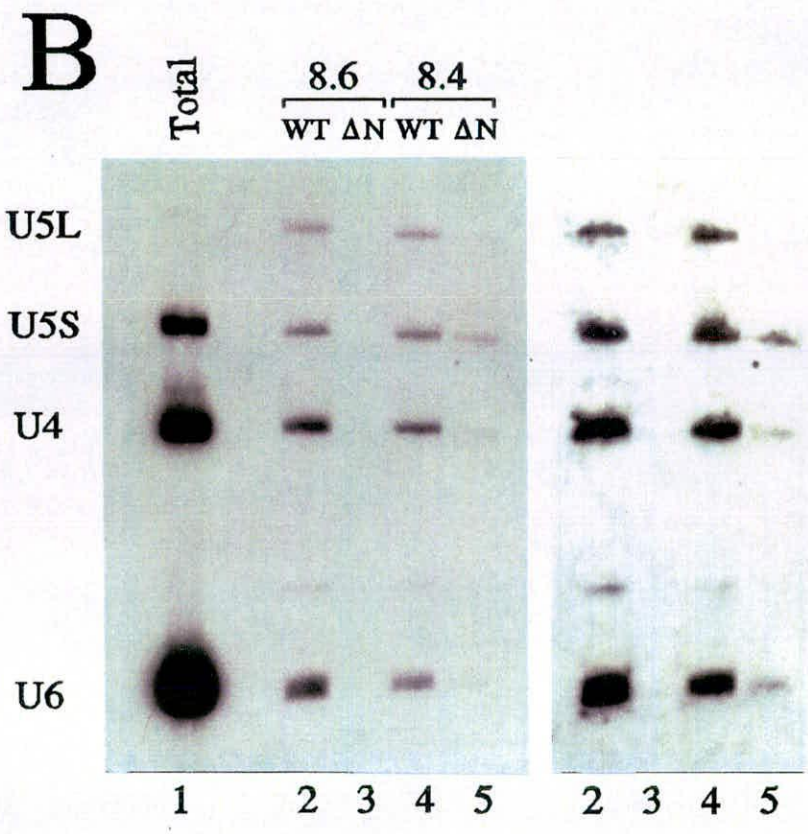
(A): Western Blot.

Splicing extract was prepared from S150-2B (grown in YPDA; WT) and AJY8.07 (grown in YPGRA; ΔN). Equal volumes of extract were fractionated by SDS-PAGE through a 6% gel. Proteins were electroblotted to Immobilon-P membrane and the Prp8p species visualised with anti-8.4 (lanes 1 and 2) and anti-8.6 antibodies (lanes 3 and 4). Positions of protein molecular weight standards (M; mass in kDa) are indicated.

(B): Immunoprecipitation.

Mock splicing reactions assembled with S150-2B wild-type (WT) or AJY8.07 (ΔN) splicing extract were incubated with either anti-8.4 (lanes 2 and 3) or anti-8.6 (lanes 4 and 5) anti-Prp8p antibodies immobilised on protein A-Sepharose. RNA extracted from the immunoprecipitates and 10μg of yeast total RNA (lane 1) was analysed for U4, U5 and U6 snRNAs by Northern blotting as described in Figure 4.1. Even in a longer exposure (right hand panel), snRNAs are not detected in the anti-8.6/AJY8.07 precipitate (lane 3).


A
8.4
8.6
M
WT ΔN WT ΔN
205
180
116
Prp8p
1 2 3 4


B
Total
8.6
8.4
WT ΔN WT ΔN
U5L
U5S
U4
U6
1 2 3 4 5 2 3 4 5

reacts with anti-8.4 but not anti-8.6 antibodies (Figure 4.4A). Thus truncated Prp8p was detectable in splicing complexes in AJY8.07, consistent with pJDY13 being absent from this strain.

The growth defect of AJY8.07 was analysed in detail in liquid medium (Figure 4.3). At 23°C, AJY8.07 grows in YPGRA with a doubling time of 7.5hr, 45% the rate of growth of the parent strain (JDY8.05 doubling time: 3.4hr). After transfer from 23°C to 34°C, cells continue to divide but deviate from logarithmic growth after about six hours, and division ceases after 9 hours. To determine if temperature-sensitivity is associated with a splicing defect, cells were harvested after 6 hours growth, and total RNA prepared. Northern blots were hybridised to a labelled fragment of the intron-containing *RP28* gene (Figure 4.2C), which under normal circumstances is efficiently spliced. In RNA prepared from JDY8.05 grown at 23°C and 34°C, *RP28* mRNA was present in far greater abundance than pre-mRNA, demonstrating that splicing is active at both temperatures. Transfer of JDY8.05 to glucose causes depletion of Prp8p (Brown and Beggs, 1992), and as expected pre-mRNA accumulated relative to mRNA. Similarly, in AJY8.07 pre-mRNA accumulates at the restrictive temperature (34°C); therefore the temperature-sensitive growth defect of this strain is associated with a splicing defect.

4.5 Disruption/Truncation of the Chromosomal *PRP8* Locus.

There are several possible explanations for the temperature sensitivity of strain AJY8.07. Truncated Prp8p may be more labile than the wild-type version, alternatively the effect of temperature may be indirect, affecting the process of splicing itself such that truncated Prp8p no longer can associate at the higher temperature (for example). It must also be considered possible that temperature affects expression from the GAL1 promoter in this strain. To distinguish these possibilities, a strain with a similar genotype to AJY8.07 (*GAL1:prp8ΔN*) was constructed, but in a different genetic background. Reproduction of the splicing defect at higher temperatures in this strain would support the hypothesis that the truncated protein was intrinsically sensitive to heat-inactivation.

Figure 4.5. Disruption/Truncation of the *PRP8* Chromosomal Locus.

A: Structure of the *PRP8* chromosomal locus in S150-2B, and of the fragment for targeted integration.

The *Xba*I and *Eco*RI sites upstream of the *PRP8* locus were mapped by Southern blotting (Brown, 1992) and their positions are shown approximately. The fragment for integration consisted of the *PRP8* upstream region (hatched box), the *URA3* gene, and the GAL1 promoter fused to the *Xba*I-*Eco*RI region of the *PRP8* coding sequence. These DNA fragments were assembled in pBluescript to create pAJ11 (see below). S150-2B cells were transformed with roughly 1μg of the *Sal*I to *Eag*I fragment from this plasmid, and Ura^+ colonies selected. The dashed lines indicate regions of homology to direct recombination events to the *PRP8* locus. Scale is indicated by the bar (1kb). Key for restriction enzymes: X=*Xba*I, N=*Nhe*I, S=*Sal*I, R=*Eco*RI, E=*Eag*I, N/X=fusion of *Nhe*I and *Xba*I sites, (N)=position of hexamer mutagenised to create *Nhe*I site.

B: Southern Blot.

Genomic DNA was isolated from S150-2B and from the three isolates of the putative disruptant, AJY8.09. After restriction digestion with *Eco*RI (left panel) or *Xba*I (right panel), DNA was fractionated by electrophoresis through an 0.8% agarose gel and blotted to Hybond-N. Membranes were hybridised sequentially to the DNA fragments indicated in Figure 4.5A above (probes 1 and 2) ^{32}P-labelled by the random priming method. Markers (M) were λ DNA fragments from a *Bst*EII digest. Probe 2 is specific for the N-terminal region of *PRP8* which is expected to be absent from *prp8ΔN* in AJY8.09.

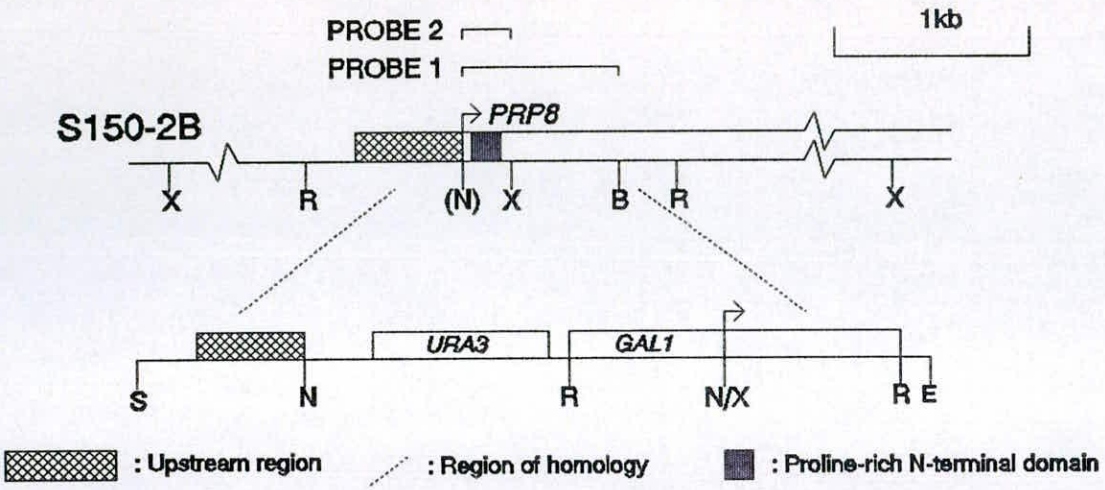
Construction of pAJ11.

Step 1: From pJDY6, the *Nhe*I fragment was removed and the remainder religated.

Step 2: The *Sal*I to *Eco*RI fragment (containing *PRP8* upstream sequence and *URA3*) was subcloned to pBluescript KS^-.

Step 3: The *EcoRI* fragment from pAJY9 (containing the *GAL1* promoter and the N-terminal end of *prp8ΔN*) was inserted, creating pAJ11.

A


PROBE 2
PROBE 1
1kb
S150-2B
PRP8
X
R
(N)
X
B
R
X
URA3
GAL1
S
N
R
N/X
R
E
: Upstream region
: Region of homology
: Proline-rich N-terminal domain

B

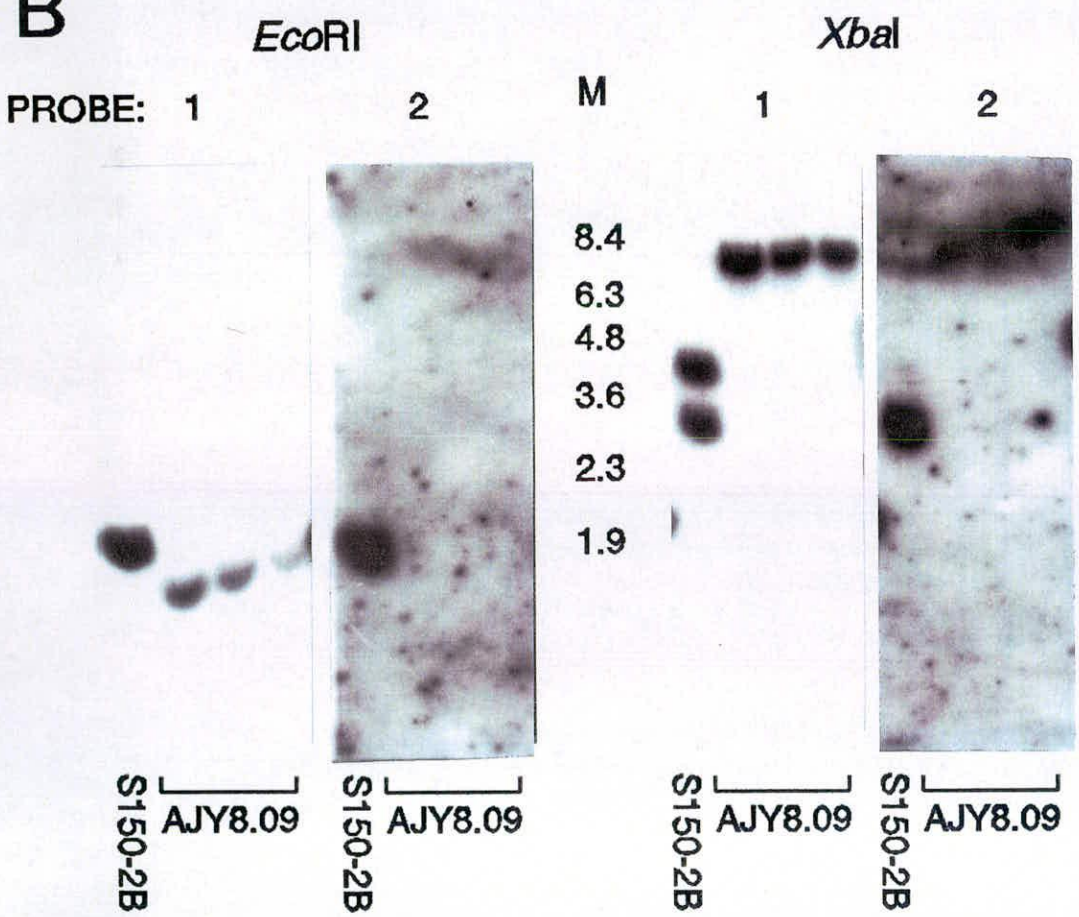

EcoRI
XbaI
PROBE: 1
2
M
1
2
8.4
6.3
4.8
3.6
2.3
1.9
S150-2B
AJY8.09
S150-2B
AJY8.09
S150-2B
AJY8.09
S150-2B
AJY8.09

Figure 4.6: Analysis of the Growth Rate of AJY8.09.

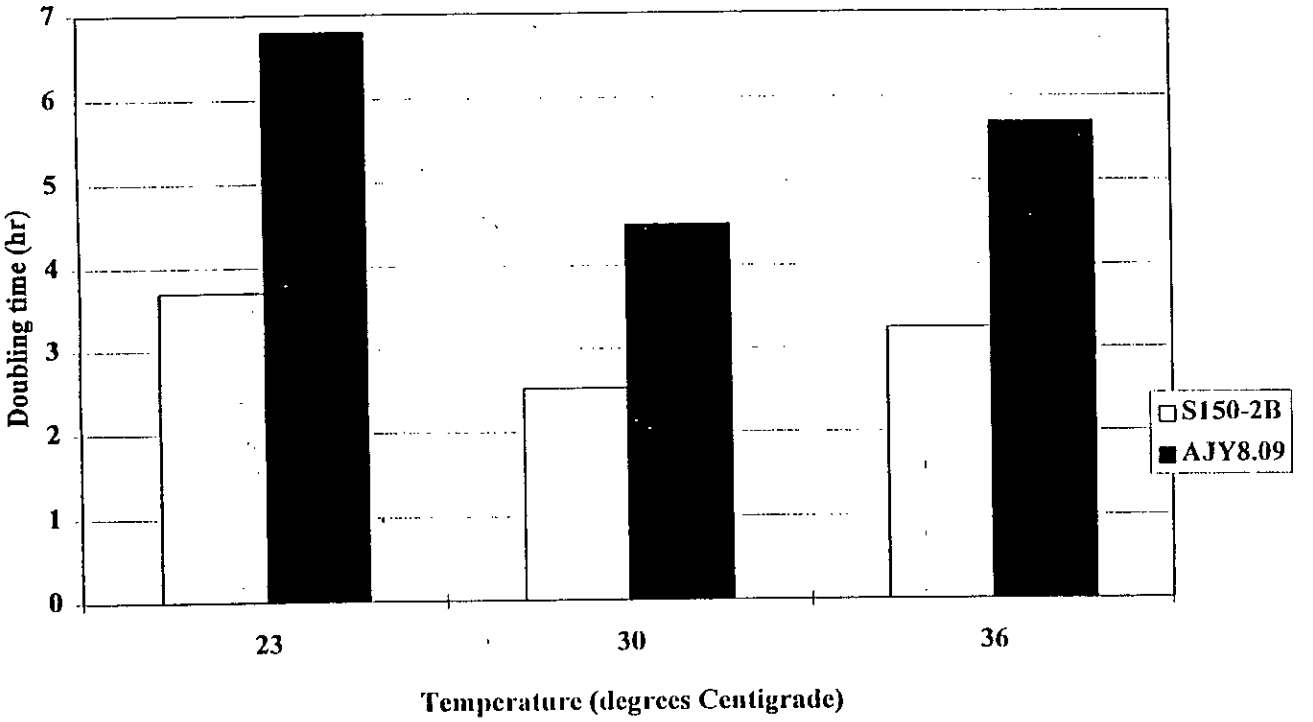
(A): AJY8.09 Growth rate.

Mid-log cultures (YPGRA) of AJY8.09 were heat-shocked to 30°C or 34°C by dilution into pre-warmed medium and OD_{600} monitored as before (Figure 4.3). No deviation from logarithmic growth could be detected over 40hrs, consistent with the ability of this strain to grow at 34°C on agar plates (not shown). The rate of logarithmic growth of this strain and of the parent strain S150-2B at 23, 30, and 34°C is presented.

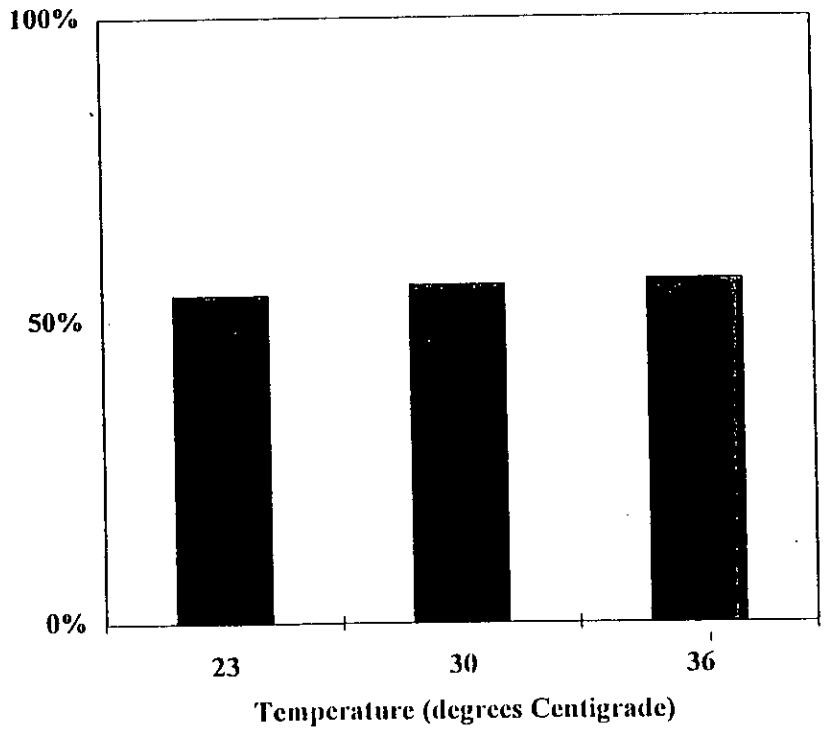
(B): AJY8.09 growth rate relative to S150-2B.

AJY8.09 growth rate as a proportion of S150-2B growth rate is presented. AJY8.09 at approximately 55% the rate of S150-2B at all three temperatures tested, and is not temperature-sensitive under these conditions.

A


7
6
5
4
3
2
1
0
Doubling time (hr)
23
30
36
Temperature (degrees Centigrade)
S150-2B
AJY8.09

B


100%
50%
0%
23
30
36
Temperature (degrees Centigrade)

The strain was constructed by integration into the chromosome. The integration event (Figure 4.5A) replaces the *PRP8* upstream region with the *URA3* marker and the *GAL1* promoter, and deletes the N-terminal domain. The DNA fragment for integration was assembled in pBluescript and removed as a linear fragment with homologous sequence at each end to direct recombination to the *PRP8* locus. Cells of the wild-type strain S150-2B (*ura3*) were transformed with this fragment, and integrants were selected on YMGRcas lacking uracil. As linear fragments are not stable in yeast, any Ura^+ colonies obtained were expected to have acquired the *URA3* gene by insertion into the genome. After prolonged incubation (14 to 21 days at 23°C) three colonies were obtained. The requirement for incubation of such length probably stemmed from a combination of three factors: i) incubation was at a non-optimal temperature on a non-optimal carbon-source, ii) the disruptant was expected to have a growth defect, and iii) a lag phase before full induction of the *GAL1* promoter. The three transformant colonies were galactose-dependent, as expected for an integration event placing an essential gene (*PRP8*) under the *GAL1* promoter. Southern blotting (Figure 4.5B) confirmed that the expected disruption had taken place. In particular one probe (probe 2) was specific to the N-terminal domain, and detected no band in genomic DNA from the Ura^+ colonies, confirming that the integration event deletes the N-terminal domain. This strain was named AJY8.09 (*ura3, prp8::URA3-GAL1:prp8ΔN*).

4.6 Phenotype of AJY8.09.

AJY8.09 has a growth defect in liquid YPGRA media, growing at 30°C at approximately 55% the rate of the parent strain S150-2B. AJY8.09 is not temperature-sensitive, growing at 23°C and 36°C at the same rate respective to the parent strain S150-2B (Figure 4.6). Only the phenotype of AJY8.07 of slow growth at 23°C is reproduced in AJY8.09, and not the inability to grow at high temperatures. Temperature sensitivity is not a general property of yeast strains which overexpress *prp8ΔN*, and truncated Prp8 protein is not intrinsically sensitive to heat inactivation. AJY8.07 temperature sensitivity is most likely a result of the JDY8.05 background (see discussion).

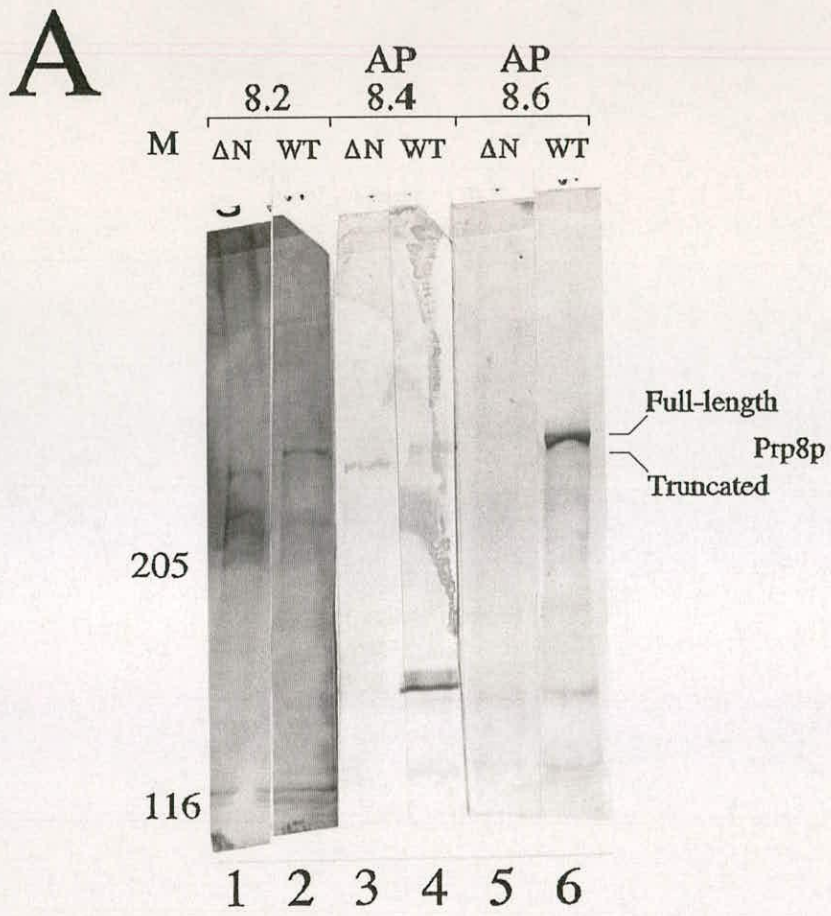
Figure 4.7: Truncated Prp8p in AJY8.09 Extracts.

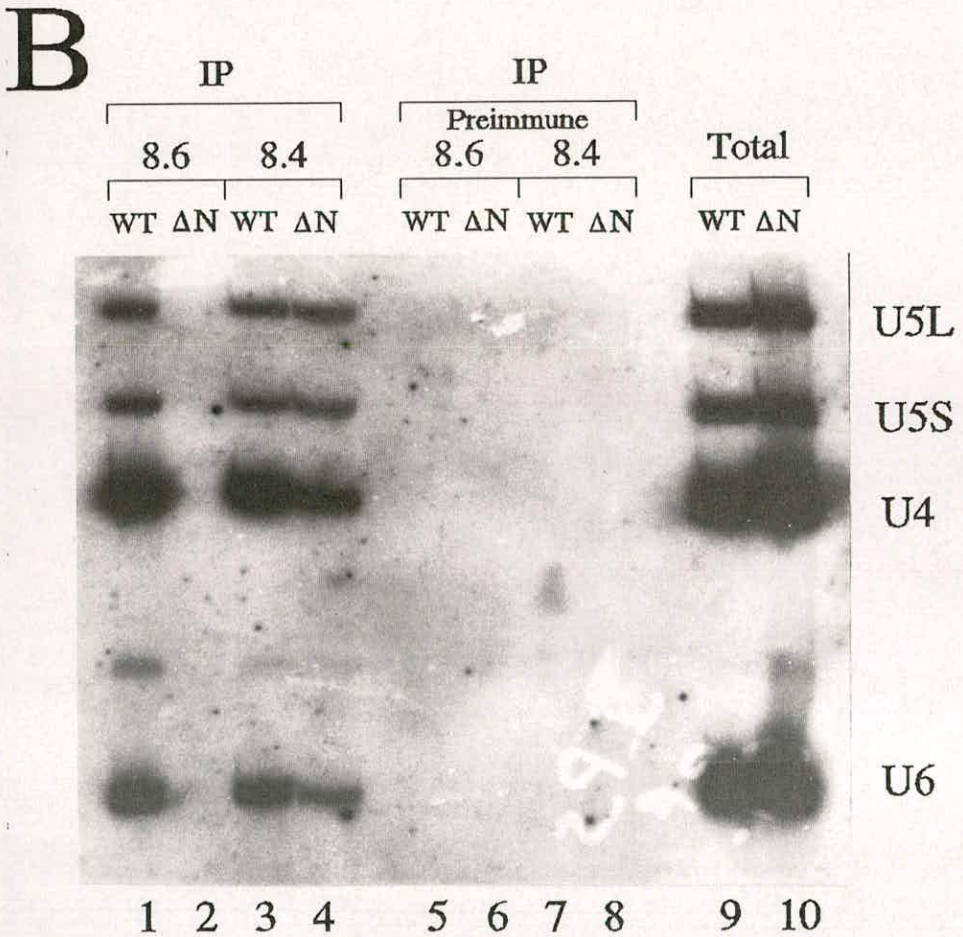
(A): AJY8.09 contains truncated Prp8p.

Splicing extracts were prepared from S150-2B (wild-type; WT) or AJY8.09 (*prp8::URA3-GAL1:prp8ΔN*; ΔN) cells grown in YPGRA (inducing) medium. Extracts (100mg protein per lane) were fractionated by SDS-PAGE through a 6% gel, electroblotted to Immobilon-P, and decorated with anti-8.2 crude serum (lanes 1 and 2), or affinity purified anti-8.4 antibodies (lanes 3 and 4), or affinity purified anti-8.6 antibodies (lanes 5 and 6). Protein molecular weight markers (M) are indicated (size in kDa).

(B): U4/U6.U5 snRNP level is reduced in the AJY8.09 extract.

Mock splicing reactions were assembled with S150-2B and AJY8.09 extracts (from (A) above; 1mg protein per mock reaction) and incubated with protein A Sepharose-coupled anti-8.6 (lanes 1 and 2) or anti-8.4 (lanes 3 and 4) immune serum, or the corresponding preimmune serum (lanes 5 to 8). SnRNAs were analysed as described in Figure 4.1B. For lanes 9 and 10, aliquots of complete extract (100mg protein) were deproteinised directly, and analysed for snRNA. 10% fractions of the supernate of each immunoprecipitation were also analysed for snRNA (not shown) to demonstrate that the precipitations had not been affected by nucleases and that full-length U4, U5 and U6 were available for coprecipitation in each case.


A
AP
AP
8.2
8.4
8.6
M
ΔN WT ΔN WT ΔN WT
Full-length
Prp8p
Truncated
205
116
1 2 3 4 5 6


B
IP
IP
Preimmune
8.6 8.4
8.6 8.4
Total
WT ΔN WT ΔN
WT ΔN WT ΔN
WT ΔN
U5L
U5S
U4
U6
1 2 3 4 5 6 7 8 9 10

Splicing extract from strains AJY8.09 and S150-2B grown at 30°C in YPGRA was prepared and analysed by Western blotting for Prp8 protein. AJY8.09 contained a Prp8p species which migrated faster than wild-type Prp8p, and which failed to react with anti-8.6 antibodies; the same pattern as AJY8.07. With anti-8.4 antibodies, the truncated Prp8p band of AJY8.09 is detected with similar, or slightly greater intensity than wild-type Prp8p in S150-2B. As shown in Chapter 3, this is roughly the protein level which is expected to result from *GAL1* induction in this system. This suggests that the truncation does not hyper-stabilise the protein. Although it cannot be concluded from this experiment that the truncation does not decrease stability, it does indicate that an effect on protein stability cannot be the only defect of *prp8ΔN*. This conclusion comes from consideration of JDY8.05 which expresses full-length Prp8p at less than 10% wild-type levels but has no growth defect (Brown and Beggs, 1992). Therefore Prp8p is normally present in excess in wild-type cells. AJY8.09 produces truncated protein at or near wild-type levels but still has a significant growth defect.

The AJY8.09 extract is functional in standard splicing assays, although much less active than an extract from S150-2B at identical concentration. Native gel electrophoresis detects complex III, the prespliceosome, but little complex I or II, consistent with the low level of splicing activity, and suggesting that up to the point where Prp8p is involved, spliceosome assembly proceeds as normal (data not presented). Immunoprecipitation from these extracts with anti-8.4 antibodies detected very similar overall levels of U5 snRNP particle in the two extracts, but from the AJY8.09 extract levels of coprecipitated U4 and U6 snRNAs were reduced by 25%, suggesting a reduced level of U4/U6.U5 tri-snRNP compared to the parent strain S150-2B. This is likely to be (at least one of) the cause(s) of the *in vitro* splicing defect in this extract, and possibly the growth defect of AJY8.09. It is consistent with a role for the repetitive, proline-rich N-terminal domain in facilitating the association between the U5 and U4/U6 snRNP particles, although the domain is not essential for this process.

4.7 Discussion.

Behaviour of the truncation mutants varies according to the presence of other mutant forms of Prp8p in cells. For ease of interpretation, results from *prp8*

temperature-sensitive backgrounds (Section 4.3) and from *prp8* null backgrounds (Sections 4.4 to 4.6) are discussed separately.

4.7.1 Null Backgrounds.

The results show that partial deletion of the N-terminal domain does not abolish domain function but causes intermediate defects. Removal of the entire N-terminal domain reduces activity to very low levels. Removal of roughly half of the N-terminal domain to give *prp8ΔP2* reduces activity to a level intermediate between the full truncation and full-length wild-type. Extrapolating, it is expected that versions of Prp8p with 1 or 3 copies of the LP_nG motif would similarly possess intermediate activity.

With both truncations, overexpression has a compensatory effect. The viability of yeast bearing *prp8ΔN* is dependent on overexpression of the allele. The strains AJY8.07 and AJY8.09 are dependent for their viability on overexpression of *prp8ΔN*. The truncated protein lacking all four copies of the LP_nG motif must still adequately carry out the functions of *PRP8*, and thus it is unlikely that catalysis in the spliceosome directly involves the Prp8p N-terminal domain. Rather, the domain acts to facilitate the splicing process, effectively reducing the required level of expression of the gene. Overexpression did not cause dominant negative phenotypes, on the contrary, overexpression compensated the deleterious effect of truncation mutation.

The discrepancy amongst the null-background strains was that the temperature-sensitive splicing defect of AJY8.07 was not reproduced in AJY8.09. At 23°C both strains had similar growth defects, growing at approximately half the rate of isogenic wild-type strains (55% for AJY8.09; 45% for AJY8.07). At higher temperatures AJY8.07 ceased to grow and had a splicing defect, whereas AJY8.09 had an identical growth rate over this temperature range relative to controls. Temperature sensitivity cannot therefore be an intrinsic property of *prp8ΔN*. A possible explanation for the discrepancy is that *prp8ΔN* is less strongly expressed in AJY8.07 at higher temperatures. Activity of the truncation mutants is critically dependent on overexpression in contrast to the wild-type which does not become limiting for growth even when expressed at low levels. It is of note that one of the two haploid strains used to construct JDY8.05 was DBY745 (Brown and Beggs, 1992), a strain suspected of

being temperature-sensitive for *GAL1* induction (J. Morran, unpublished data). This may also explain why JDY8.05 is itself temperature-sensitive (at 36°C; not presented) as it too is reliant on expression from the *GAL1* promoter for viability. Together with the findings of the previous chapter, this highlights the possibility that JDY8.05 and strains derived from it are associated with both abnormal glucose repression and abnormal galactose induction at the *GAL1* promoter. Before JDY8.05 is used in future for experiments of the type described herein, a comparative Northern analysis under a range of conditions (or alternatively a new null strain) would be helpful. For the purposes of this discussion, I assume that the temperature-sensitivity of AJY8.07 is spurious, and a consequence of the strain background, and will concentrate on AJY8.09.

AJY8.09 contains levels of the U5 snRNP identical to the parent strain S150-2B, but levels of tri-snRNP reduced by approximately 25%. This is direct evidence that a function of the N-terminal domain is to facilitate U4/U6.U5 tri-snRNP assembly. This result also implies a function earlier in the assembly process. Prp8p level is not normally limiting for assembly of the U5 snRNP (Brown and Beggs, 1992). If the protein products of *prp8ΔN* and *PRP8* had equivalent activities in the U5 snRNP assembly process (if the truncation mutation did not affect this step) then overexpression should not affect the phenotype: both AJY8.09 (*GAL1:PRPΔN*) and a strain constitutively expressing *prp8ΔN* should have similar levels of U5 snRNP and therefore similar phenotypes. Expression level markedly affects the phenotype of truncation mutations. Therefore, a potential role for the N-terminal domain is in promoting U5 snRNP assembly or stability. The action of overexpression may be to force an equilibrium in favour of the U5 snRNP.

A model with which these results are consistent is that the proline-rich N-terminal domain promotes multiple stages of spliceosome assembly, including U5 assembly and tri-snRNP assembly, and perhaps tri-snRNP association with the prespliceosome.

4.7.2 Temperature-sensitive Backgrounds.

In *prp8* temperature-sensitive strains, truncated Prp8p exhibits a greater degree of function than when in a *prp8* null strain. When *prp8ΔN* is the sole gene copy,

viability depends on overexpression, but when the same mutant is combined with temperature sensitive *prp8*, overexpression is no longer required. The *prp8-1* strain was complemented by *prp8ΔN* fully, to wild-type rates of growth, as far as can be determined by plate assays. Results from temperature-sensitive strains bearing other *prp8* alleles were consistent with this (presented in Chapter 7); thus in general the requirement for overexpression of *prp8ΔN* for viability does not apply to temperature-sensitive strains. This indicates a phenomenon of intragenic complementation. At a genetic level, the truncation and temperature-sensitive mutants interact in synergy. A possible basis for this is that the functional entity is a dimer of Prp8p, and that the heterodimer between truncated and temperature-sensitive mutants can function whereas homodimeric mutants cannot. However, immunoprecipitation detected truncated and not full-length Prp8p in splicing complexes. The stoichiometry of Prp8p within the U5 snRNP is not yet known, but if this particle is occupied by more than one Prp8p molecule, then in this situation all Prp8p species are truncated. In this instance, intragenic complementation is not evidence that the functional unit of Prp8p is a multimer. A well-studied example of intragenic complementation is that of calmodulin and its mutants. In this example, complementation is thought to occur because calmodulin, which is believed to function as a monomer, participates in numerous different processes in the cell (Ohaya and Botstein, 1994). The different calmodulin mutants can still carry out some of the essential roles of this factor, and combinations of different mutants can result in cell viability. Consequently, calmodulin mutants can be separated into intragenic complementation groups. A similar situation may pertain in the case of Prp8p, but at present there is little supporting evidence. The idea that Prp8p may have multiple independent roles in the cell is expanded in later chapters. Lastly, it is possible that truncated and temperature-sensitive proteins interact in the cell without the full-length protein entering the complexes detected in immunoprecipitation analyses. Perhaps the activity to catalyse the early steps of spliceosome assembly, or other function of the N-terminal domain, can be supplied *in trans* from a separate polypeptide to the remainder of Prp8p.

4.7.3 Truncation Mutants and Protein Stability.

Part of the impact of truncation mutation may be on protein stability. AJY8.09 contains a similar level of Prp8p as a wild-type control. According to the hypothesis that Prp8p is stable when in ribonucleoprotein complexes and unstable otherwise, much of the Prp8p in AJY8.09 is expected to be in U5 snRNP particles. Truncated Prp8p may have identical intrinsic stability to full-length, but a slower rate of assembly into the (stabilised) U5 snRNP complex, and thus simply greater access to proteolytic action. Alternatively, the truncated protein may be intrinsically more sensitive to proteolysis than full-length Prp8p and more rapidly degraded as it is synthesised, decreasing the time-space in which U5 assembly can occur. As both scenarios are predicted to be rescued by overexpression, they cannot be differentiated on the basis of these results.

Chapter 5.

A Synthetic, Functional Prp8p N-terminal Domain.

5.1 Introduction.

In a survey of the SwissProt database for proteins containing stretches of at least 10 prolines, Gerber *et al.* (1994) found four-fifths of such proteins to be known or suspected transcription factors. Proline-rich regions in transcription factors in general form protein:protein interactions with other components of the transcription initiation complex to activate transcription (for a review of transcription factors, see Mitchell and Tjian, 1989). What these other components are is gradually becoming clear (Sheldon and Reinberg, 1995). Another similarity between the N-terminal domain of Prp8p and transactivation domains is that these regions exhibit little requirement for specific sequence. This property of Prp8p is evidenced by the demonstration that partial deletion results in partial loss of function (Chapter 4). In the case of transactivation domains, functional interchangeability has been demonstrated by domain swap experiments, and is a general feature. Some activation domains have a modular structure, in so far as duplicated sub-domains can substitute for the entire domain (Seipel *et al.*, 1992). Gerber *et al.* (1994) were able to create a functional synthetic transcription factor simply by fusing the Gal4p DNA binding domain to a polyproline (or polyglutamine) homopolymer, showing that this minimal element possessed activation activity. The optimal length of homopolymer was 10 prolines (or 10 to 30 glutamines). Other polypeptides such as poly-E, poly-D or poly-QQL were inactive.

To determine the sequence requirement for a functional Prp8p N-terminal domain, chimeric *PRP8* genes were created by fusion of novel DNA fragments to the 5' end of *prp8ΔN*. Gene construction starting from existing *PRP8* constructs would have involved multiple cloning steps for each chimera. To avoid this pAJY22 was constructed, in which fusions can be made in a single step. Chimeras were expressed *in vivo* to test for a phenotype of restored wild-type function. The test utilised the *prp8* temperature-sensitive strain L149-7B (*dbf3-1*); the nomenclature of this mutant differs because *dbf3-1* was identified in a screen for mutants with a **dumbbell**-forming

phenotype (covered more fully in Chapter 6), but the mutation responsible for this phenotype is a single nucleotide change in *PRP8*. The feature which is relevant here is that *prp8ΔN* does not support growth of L149-7B at the non-permissive temperature (the reasons for this are explored in Chapter 7). Thus chimeras created from *prp8ΔN* which complement L149-7B have regained activity of the *PRP8* N-terminal domain.

5.2 Construction of pAJY22.

The plasmid pAJY22 expresses *prp8ΔN* from the *GAL1* promoter. Its construction is presented in Figure 5.1. Deletion of the N-terminal domain was achieved not by *Xba*I/*Nhe*I fusion as before, but instead by using a linker to fuse the *GAL1* promoter to the *Xba*I site of *PRP8*, enabling a novel restriction site to be included. The linker encodes the first four Prp8p codons, followed by an *Eag*I site. This site is unique in pAJY22 and is located such that fragments inserted into this site are translated to yield chimeric proteins with novel N-termini encoded by the insert.

5.3 Characterisation of pAJY22.

The truncated *PRP8* gene in pAJY22 and *prp8ΔN* in previously tested constructs (pAJY9 and pJDY7) are different, due to the addition of the extra restriction site. pAJY22 encodes the N-terminal sequence MS**GLAA**GLDELETKAE...etc, longer than the previous truncated protein (MSGLDELETKAE...etc) by 4 amino acids (additional amino acids in bold). To confirm that the addition of these four amino acids did not change the activity of the protein, pAJY22 was introduced by transformation into two temperature-sensitive *prp8* strains: SPJ8.31 (*prp8-1*) and L149-7B (*dbf3-1*). The plasmid conferred the ability to grow at the restrictive temperature (inducing conditions) to SPJ8.31, but not to L149-7B (Figures 5.3 and 5.4). This is the same result as was found with pAJY9 (*CEN, URA3, GAL:prp8ΔN*) and pJDY7 (2μ, *URA3, prp8ΔN*), and shows that *prp8ΔN* is not sufficient to complement L149-7B, for which a

Figure 5.1: Construction of pAJY22.

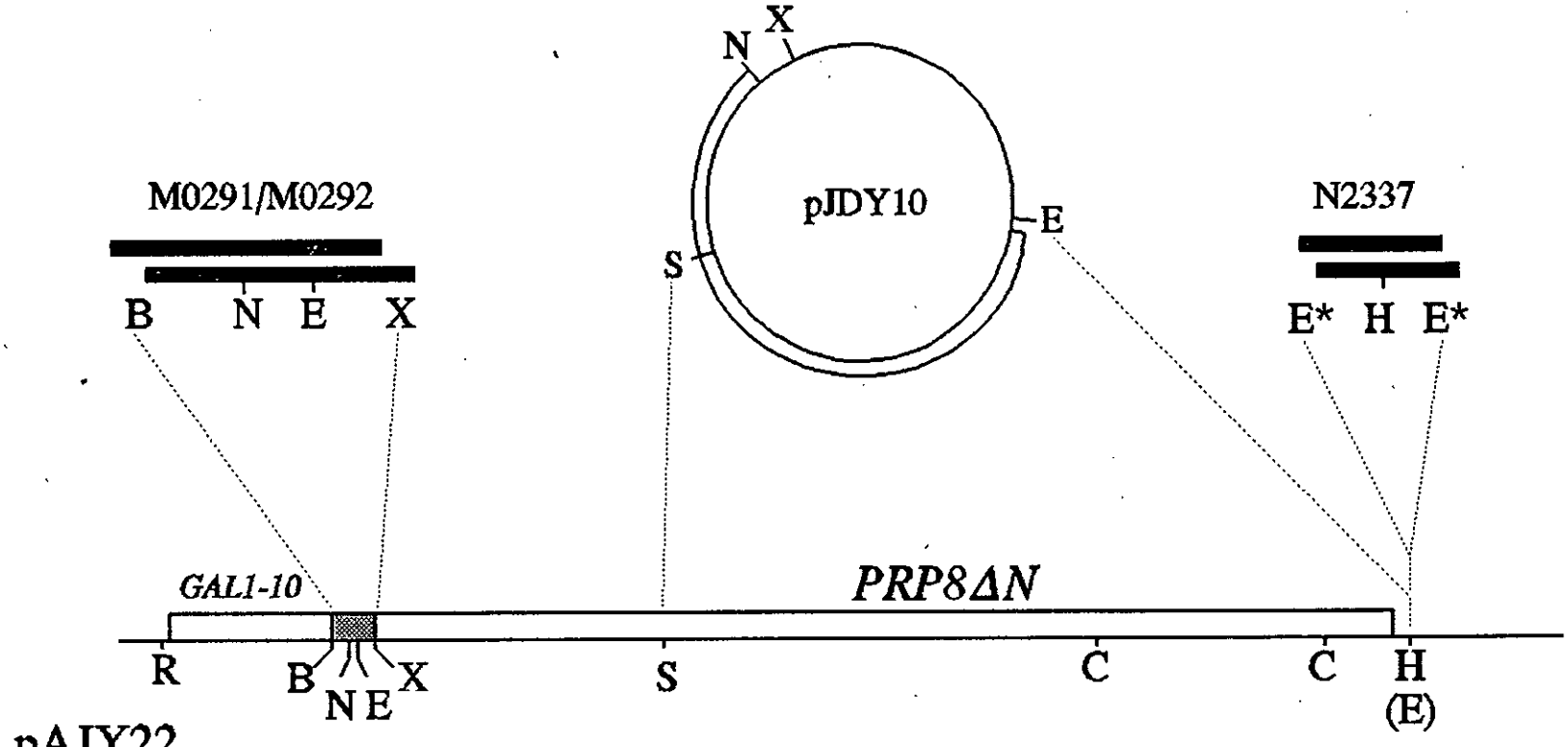
A plasmid containing a unique restriction enzyme (*Eag*I) site immediately after the start codon of *prp8ΔN* was constructed in three steps. In each step, non-unique restriction sites were used in cloning, and fragments of appropriate length were purified from partial restriction reactions. In the final construct, pAJY22, the *Eag*I site, the initial codons of *prp8ΔN*, and a consensus sequence for translational initiation are derived from the linker. The *Eag*I site was chosen because the amino acids it encodes (alanines in this reading frame) are not charged. Plasmid pAJY22 was sequenced through the M0291/M0292 linker region to confirm monomeric insertion. pAJY22 is identical to the previous construct pAJY9 except for the 12 extra nucleotides within *prp8ΔN*, and the palindromic linker inserted into the vector (pBM125) *Eag*I site.

Step 1: The *Bam*HI-*Xba*I fragment of pAJY6 was replaced with the linker M0291/M0292, to create pAJY17.

Step 2: The remainder of the *PRP8* gene was introduced into pAJY17 as a *Sal*I-*Eag*I fragment, to create pAJY21.

Step 3: The *Eag*I site deriving from pBM125 (downstream of *prp8ΔN*) was destroyed by insertion of the palindromic N2437 linker.

Key. B: *Bam*HI, N: *Nhe*I, E: *Eag*I, R: *Eco*RI, X: *Xba*I, S: *Sal*I, H: *Hind*III, C: *Cla*I, (E): position of *Eag*I site in pBM125. E^*: Complementary to *Eag*I end; cannot be cleaved by *Eag*I after ligation.



pAJY22

Vector: pBM125

Parent: pAJY6

Figure 5.2: Insertion of Linkers into the pAJY22 *Eag*I Site.

pAJY22 was linearised at the unique *Eag*I site, and treated with phosphatase. Complementary oligodeoxynucleotides were treated with T4 polynucleotide kinase and annealed to form linkers with *Eag*I-complementary ends, and introduced into pAJY22 by ligation. Non-palindromic linkers were designed with so that a site for the 8-base recognition *Not*I enzyme (which overlaps with an *Eag*I site) was generated by insertion in only one of the two possible orientations, facilitating identification of recombinants. Insert length was determined by PCR (using primers G0879 and A153), and inserts were sequenced. pAJY32 was constructed similarly, by insertion of the CR1/CR1 linker into pAJY27.

pAJY29 was a common product of the ligation of N6930/N6931 into pAJY22, although its sequence is shorter than expected. From its sequence, the linker in pAJY29 appears to be truncation of the N6930/N6931 linker missing a central region. These oligodeoxynucleotides were not purified by HPLC and the unexpectedly short linker in pAJY29 may be due to a shortened oligodeoxynucleotide from aberrant synthesis. As the linker in pAJY29 maintained the reading frame, it was introduced to L149-7B with the other plasmids of the series, and tested.

Key for restriction enzymes: B, *Bam*HI; N, *Nhe*I; E, *Eag*I; X, *Xba*I.

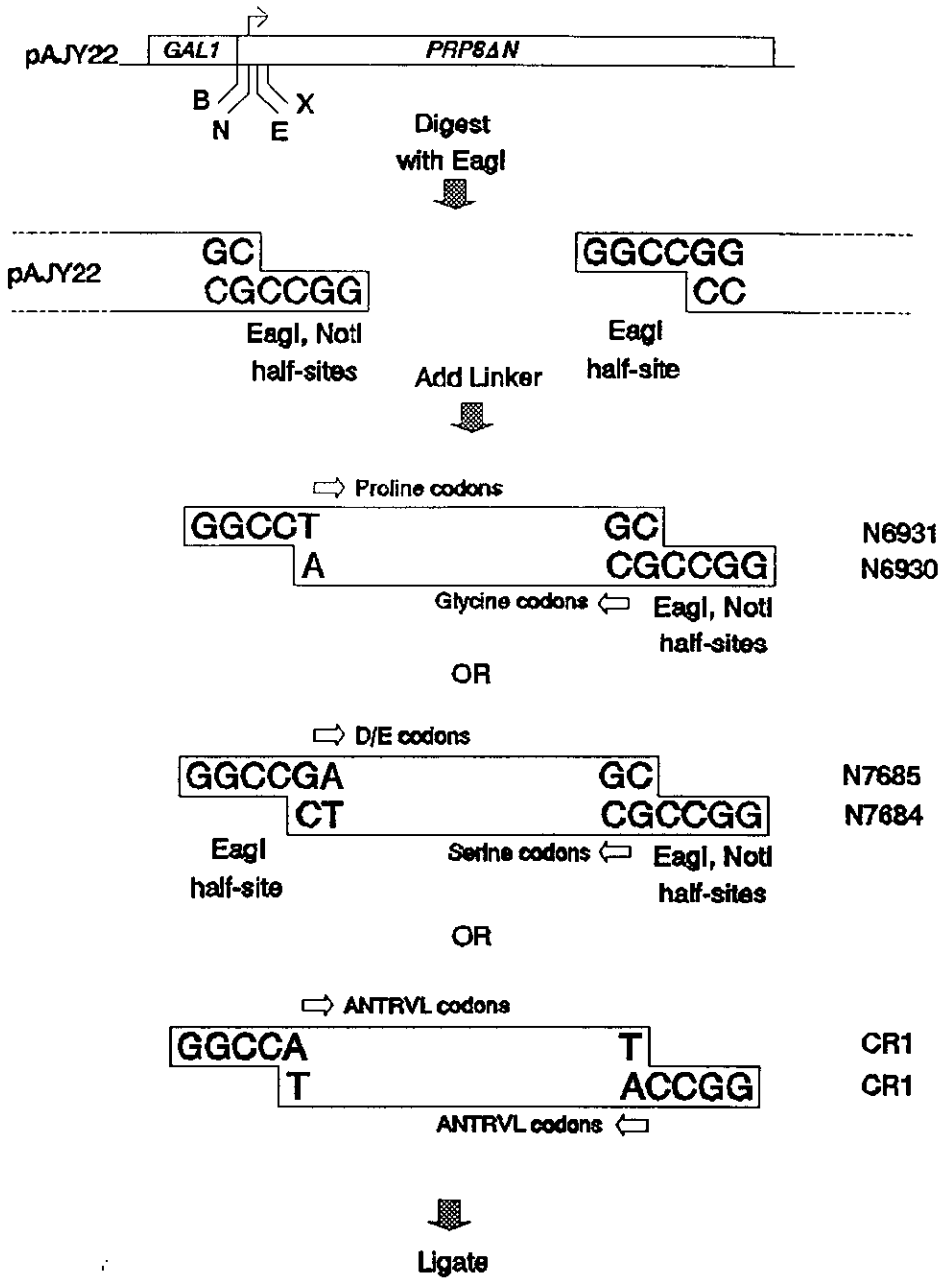

pAJY22
GAL1
PRP8ΔN
B
X
N
E
Digest
with EagI
pAJY22
GC
CGCCGG
EagI, NotI
half-sites
GGCCGG
CC
EagI
half-site
Add Linker
Proline codons
GGCCT
A
GC
CGCCGG
N6931
N6930
Glycine codons
EagI, NotI
half-sites
OR
D/E codons
GGCCGA
CT
GC
CGCCGG
N7685
N7684
EagI
half-site
Serine codons
EagI, NotI
half-sites
OR
ANTRVL codons
GGCCA
T
T
ACCGG
CR1
CR1
ANTRVL codons
Ligate

Figure 5.2

functional N-terminal domain is required. The tetrapeptide GLAA extension of pAJY22 is not sufficient to restore wild-type function. For the purposes of this investigation, the two versions of *prp8ΔN* were regarded as equivalent.

5.4 Generation of Novel Prp8p N-termini.

Genes expressing chimeric proteins, with novel Prp8p N-termini, were generated by inserting linkers into the *Eag*I site of pAJY22 (Figure 5.2). Linkers were designed to regenerate a unique restriction site, so that a synthetic N-terminal domain could be built up from consecutive modules. The linkers used in this study were formed from complementary oligodeoxynucleotide pairs N6930/N6931 and N7584/N7585, and the self-complementary oligodeoxynucleotide CR1. The N6930/N6931 linker encodes the $ALP_{10}GAA$ pentadecapeptide (15-mer), or $A_3P(G)_{10}K$ in the opposite orientation. The N7584/N7585 linker encodes the $ADE_2(DE)_3ELA$ tridecapeptide (13-mer), or A_2S_{11} in the opposite orientation. The CR1 linker encodes the ANTRVL hexamer. The peptide sequence of each of the chimeric *PRP8* genes constructed and tested in this investigation is presented in Table 5.1.

Table 5.1

Predicted peptide sequence from the start codon to the first acidic residues (bold; equivalent to amino acids 81, 82 and 84 of wild-type Prp8p, see Figure 3.1) is shown.

Plasmid.	Insertion.	N-terminal Sequence.
pAJY9		MSGL**DELE**...
pAJY22		MSGLAAGL**DELE**...
pAJY23	N7685/N7684	MSGLAAASSSSSSSSSSSAGL**DELE**...
pAJY24	N7685/N7684	MSGLAADEEDEDEDEELAAGL**DELE**...
pAJY25	N7685/N7684 dimer	MSGLAADEEDEDEDEELAADEEDEDEDEELAAGL**DELE**...
pAJY27	N6930/N6931	MSGLAALPPPPPPPPPPGAAAGL**DELE**...
pAJY28	N6930/N6931	MSGLAAAAPGGGGGGGGGGKASL**DELE**...
pAJY29	See legend, Fig. 5.2	MSGLAALPPPPGAAAGL**DELE**...
pAJY30	CR1/CR1	MSGLAANTRVLAGL**DELE**...
pAJY31	CR1/CR1 dimer	MSGLAANTRVLANTRVLAGL**DELE**...
pAJY32	N6930/N6931 and CR1/CR1	MSGLAALPPPPPPPPPPGAAANTRVLAGL**DELE**...

5.5 Complementation of L149-7B (*dbf3-1*) by Chimeras.

As indicated above, complementation of the *dbf3-1* strain L149-7B provides a rapid test for function of the N-terminal region because in this strain Prp8p requires the N-terminal domain for complementation. Any chimeric *PRP8* gene which complements this strain has gained a function equivalent to that of the N-terminal

domain. Strain SPJ8.31 is complemented irrespective of an N-terminal domain (by *PRP8* or by *prp8ΔN*), and complementation of this strain was a test of expression. The first series of plasmid constructs (pAJY23, 24, 25, 27, 28 and 29) all complemented SPJ8.31, demonstrating that in each case *PRP8* activity was expressed (Figure 5.3). Plasmids pAJY23, 24, 25 and 28 failed to complement *dbf3-1*, showing that the chimeras did not confer wild-type *PRP8* function (Figure 5.4). Either polyglycine, polyserine and polyacid moieties at the N-terminus are non-functional or, possibly, the chimeras are unstable and readily cleaved by proteolysis. In contrast pAJY27, which encodes Prp8p with one copy of LP_nG (n=10), supports L149-7B growth at 37°C on galactose, demonstrating that a level of N-terminal domain function has been restored. pAJY29, which encodes the same protein as pAJY27 except with a shorter proline homopolymer (n=4), did not complement L149-7B. Therefore the minimum polyproline length for complementation is greater than four. These data support the notion that the proline residues are of key importance in the Prp8p N-terminal domain.

5.6 The Next Generation of Chimeras.

In the protein expressed from pAJY27, the LP_nG motif is followed by acidic residues (see Table 5.1). Likewise in wild-type Prp8p each LP_nG motif is followed by a cluster of acidic residues. In the case of pAJY27, six uncharged residues separate LP_nG and the first acidic residue; in wild-type the separation is 2 or 3 residues. To determine if the juxtaposition of LP_nG and an acidic cluster is required for function, the CR1/CR1 linker was inserted into the pAJY27 *Eag*I site, to create plasmid pAJY32 (see Table 5.1). pAJY32 complements L149-7B (Figure 5.4), indicating that the LP_nG motif has no requirement for following acidic residues and can function even when followed by a basic residue. As a control, the CR1/CR1 linker was inserted into pAJY22, creating pAJY30 (monomer insert) and pAJY31 (dimer insert); these plasmids were unable to complement L149-7B. These data confirm that the LP_nG motifs, and not the overall acidic charge of the N-terminal domain are the most important sequence feature.

Figure 5.3: Galactose-dependent Expression of *PRP8* Activity from Plasmids of the Series pAJY22 to pAJY32.

Cells of the temperature-sensitive strain SPJ8.31 (*prp8-1*) harbouring various *URA3* plasmids as shown were streaked to selective, minimal plates for growth under repressing (YMGlucas, glucose) or inducing (YMGRcas, galactose) conditions. Plates were incubated for 6 days at 23°C or 36°C. Plasmid-borne *PRP8* alleles were under the transcriptional control of the *GAL1* promoter. The nature of the amino acid sequence encoded by the *PRP8* alleles of these plasmids is given below, and in more detail in Table 5.1.

pBM125	: vector.
pBM-PRP8	: Prp8 (wild-type).
pAJY9	: Prp8ΔN.
pAJY22	: Prp8ΔN.
pAJY23	: polyserine+Prp8ΔN.
pAJY24	: polyacid+Prp8ΔN.
pAJY25	: $(polyacid)_2$+Prp8ΔN.
pAJY27	: polyproline+Prp8ΔN, n=10.
pAJY28	: polyglycine+Prp8ΔN.
pAJY29	: polyproline+Prp8ΔN, n=4.
pAJY30	: linker+Prp8ΔN.
pAJY31	: $(linker)_2$+Prp8ΔN.
pAJY32	: polyproline+linker+Prp8ΔN.

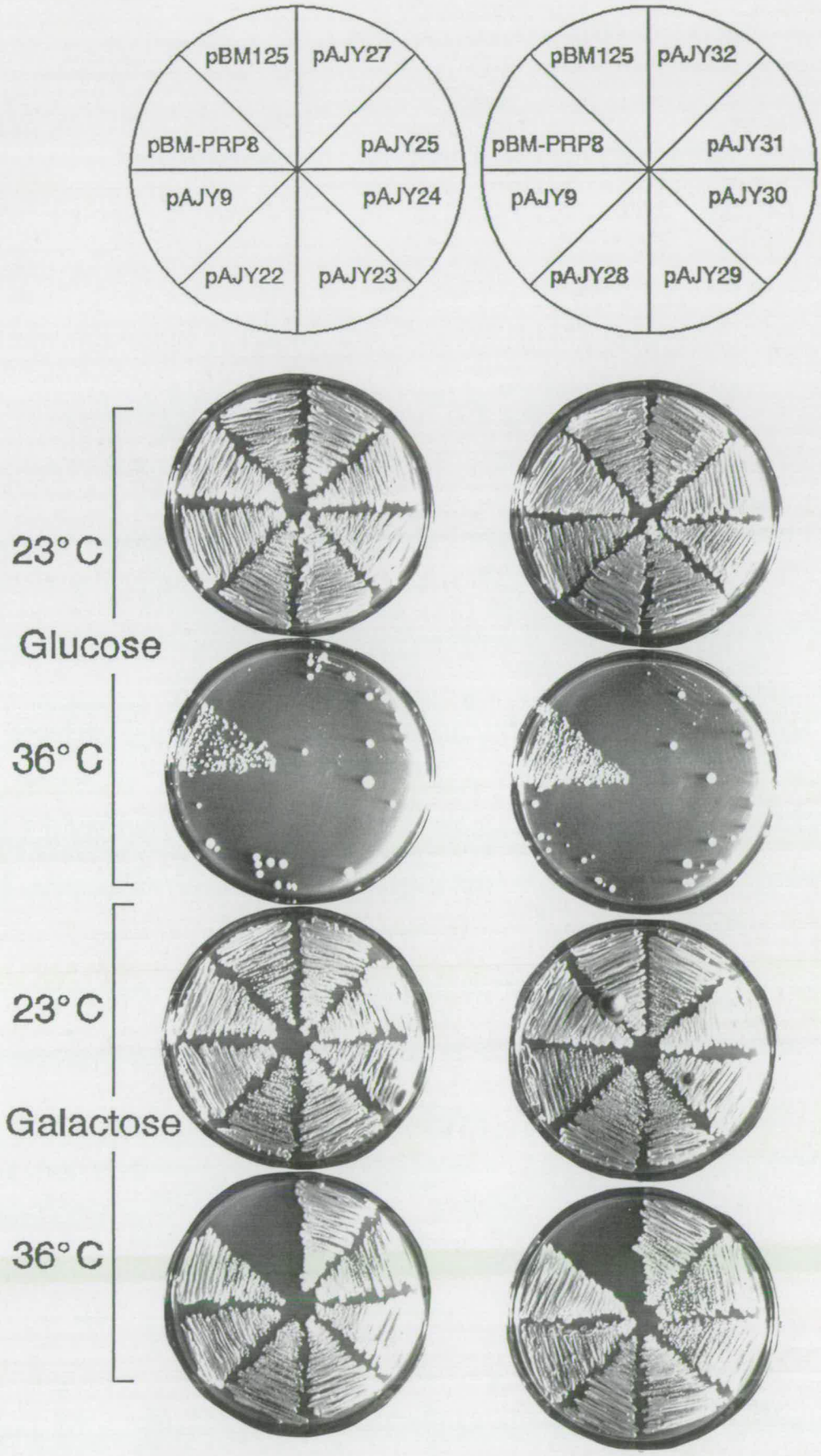

pBM125
pAJY27
pBM-PRP8
pAJY25
pAJY9
pAJY24
pAJY22
pAJY23
pBM125
pAJY32
pBM-PRP8
pAJY31
pAJY9
pAJY30
pAJY28
pAJY29
23° C
Glucose
36° C
23° C
Galactose
36° C

Figure 5.4: Complementation of L149-7B (*dbf3-1*) by Plasmids of the Series pAJY22 to pAJY32.

Following single colony purification, transformant cells of the *prp8* temperature sensitive-strain L149-7B (*dbf3-1*) were streaked to selective, minimal plates for growth under inducing conditions (YMGRcas). Plates were incubated for 5 days at 30°C or 37°C. The nature of the amino acid sequence encoded by the chimeric *PRP8* genes in plasmids borne by these cells is given below. All plasmid-borne *PRP8* alleles were under the transcriptional control of the *GAL1* promoter. For the precise peptide sequence, see Table 5.1.

pBM125 : vector.

pBM-PRP8 : Prp8 (wild-type).

pAJY9 : Prp8ΔN.

pAJY22 : Prp8ΔN.

pAJY23 : polyserine+Prp8ΔN.

pAJY24 : polyacid+Prp8ΔN.

pAJY25 : $(polyacid)_2$+Prp8ΔN.

pAJY27 : polyproline+Prp8ΔN, n=10.

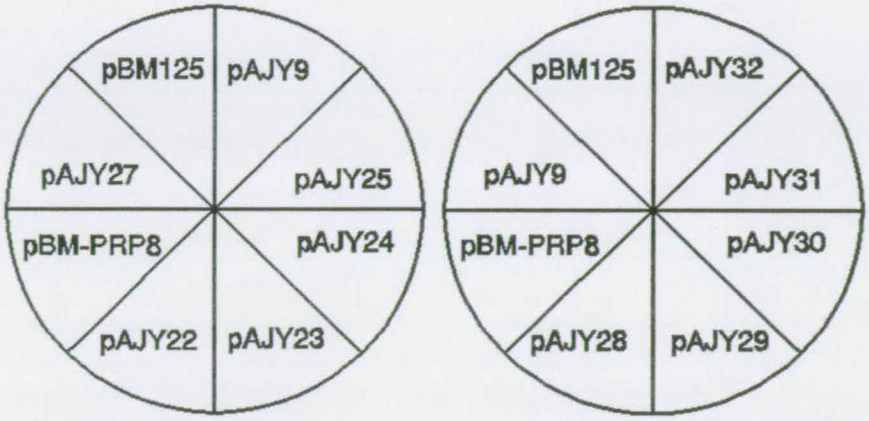
pAJY28 : polyglycine+Prp8ΔN.

pAJY29 : polyproline+Prp8ΔN, n=4.

pAJY30 : linker+Prp8ΔN.

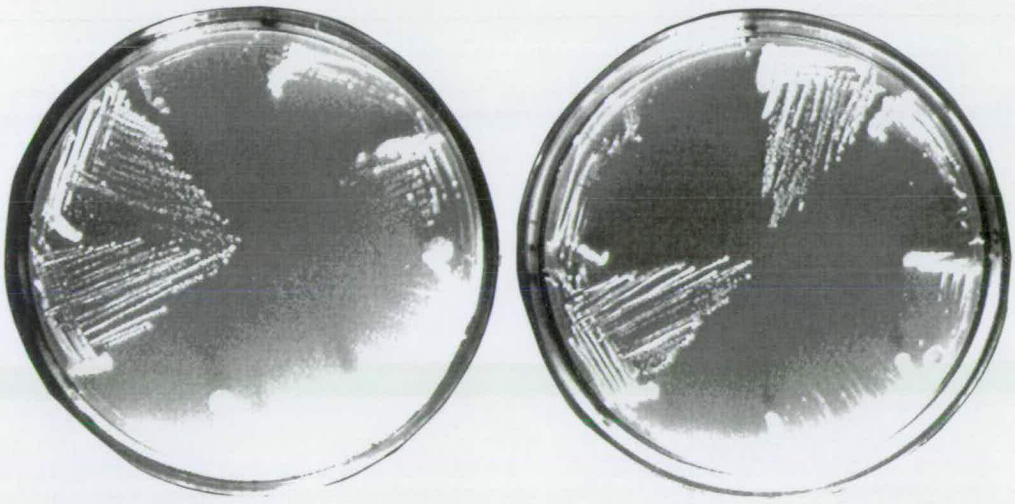
pAJY31 : $(linker)_2$+Prp8ΔN.

pAJY32 : polyproline+linker+Prp8ΔN.


pBM125
pAJY9
pAJY27
pAJY25
pBM-PRP8
pAJY24
pAJY22
pAJY23
pBM125
pAJY32
pAJY9
pAJY31
pBM-PRP8
pAJY30
pAJY28
pAJY29

30° C

37° C



5.7 Discussion.

5.7.1 Synthetic N-terminal Domains.

This chapter describes a preliminary attempt to determine the minimum sequence required at the N-terminus of Prp8p to confer function. The N-terminal domain of wild-type Prp8p was replaced by novel peptide sequence, and the resultant chimeras tested for complementation of the *prp8* strain L149-7B (*dbf3-1*). This rapid test of function is possible from the empirical observation that *prp8ΔN* does not itself complement this strain, in contrast with other *prp8* strains. Instead, a functional N-terminal domain is required (ie *prp8ΔP2* does complement L149-7B). This is due to an aspect of the strain background, and is discussed more fully in Chapter 7. A limitation of the test is that it is non-quantitative: the chimeras with restored activity may only have a small fraction of wild-type function, and non-complementing constructs may still represent improvements upon the *prp8ΔN* full deletion. This line of enquiry could be carried further in the future using a quantitative assay of splicing such as that based on expression of an intron-containing *CUP1* gene to confer resistance to copper (Lesser and Guthrie, 1993). If necessary, transfer of the chimeric genes generated in this study to the control of the endogenous *PRP8* promoter (for example in pAJY12) could be achieved in a single cloning step using the *Nhe*I restriction site in pAJY22.

These results show that the Prp8p N-terminal domain has no absolute functional requirement for acidity. The minimal functional sequence element which has been identified is a stretch of ten proline residues, as the motif LP_nG. The contribution of the non-proline residues in this motif (leucine and glycine or serine) has not yet been analysed. A predicted N-terminal domain containing acidic residues but no prolines (from pAJY24 and pAJY25) did not complement L149-7B. However, the genetic test for expression (complementation of *prp8-1*) does not exclude the possibility that chimeras are sensitive to proteolytic events which remove very specifically the novel N-terminus, to leave effectively the same polypeptide as the product of *prp8ΔN*. This eventuality is unlikely, but cannot be formally discounted, and before conclusions about a purely acidic N-terminus can be made, Western blots or other demonstration of expression of the entire chimera are required.

5.7.2 Comparison of the N-terminal Domain with Other Proline-rich Regions.

This investigation, although less extensive than that of Gerber *et al.* (1994), still has revealed remarkable similarities. Fusion of the Gal4p DNA binding domain to a homopolymer of six prolines stimulated transcription of a reporter construct *in vivo* (in co-transfected HeLa cells; Gerber *et al.*, 1994). Homopolymers of ten prolines were more effective; longer proline homopolymers reduced transactivation activity. At the N-terminus of Prp8p, a homopolymer of ten prolines confers activity whereas four prolines do not. Longer stretches have not been tested with Prp8p. Gerber *et al.* (1994) also found polyglutamine to be effective, although not $(QQS)_n$ or $(QQL)_n$ peptides. It appears that glutamine has in common with proline a propensity to adopt the polyproline II helix conformation: glutamine is the second most common amino acid found in this conformation (after proline) in known crystal structures (Adzhubei and Sternberg, 1993). Proline-rich domains are often also rich in glutamine (eg in the splicing factor PSF; Patton *et al.*, 1991; Gil *et al.*, 1991), and glutamine-rich regions of transcription factors have comparable properties to proline-rich transactivation domains. The ability of glutamine at the N-terminus of Prp8p to restore activity has not yet been tested.

In this study it has not been possible to demonstrate restoration of N-terminal domain activity by peptides predicted to be strongly acidic (poly-D/E). Similarly Gerber *et al.* (1994) also found homopolymers of acidic residues (poly-D and poly-E) did not activate transcription when fused to DNA binding domains, nor did $(ELQ)_n$. It seems not to be the case that random, acidic peptides possess transactivation activity, and acidic transactivation domains may have more stringent requirements for certain sequences than do proline-rich domains.

Another striking similarity between the Prp8p N-terminal domain and proline-rich regions of other proteins is with the C-terminal domain (CTD) of the large subunit of RNA polymerase II. The Pol II CTD in yeast contains around 26 copies of the heptad repeat PTSPSYS, the number of repeats being polymorphic in wild-type yeast strains. Half of this domain can be truncated (down to 13 repeats) without affecting cell viability. Further truncation (10-12 repeats) results in conditional viability and cells containing the large subunit with less than 10 repeats are not viable (Nomet *et al.*, 1987). Likewise progressive deletion of the Prp8p N-terminal domain progressively

reduced activity (Chapter 4), and there is preliminary evidence at least that conditional viability may result. The Pol II CTD is the target of phosphorylation which regulates polymerase activity. The N-terminal domain of Prp8p has several hydroxyl amino acids (serine, threonine, tyrosine) but in pAJY27 these have all been deleted except the serine at position 2 adjacent to the start methionine. As pAJY27 encodes a chimera with restored N-terminal domain activity, phosphorylation of the N-terminal domain is unlikely to have an important role in modulating Prp8p activity in an analogous way to RNA Pol II.

5.7.3 Future Experiments.

The Prp8p N-terminal domain may have a stringent requirement for contiguous prolines (as in the LP_nG motif), alternatively any peptide with overall proline-richness may function. Analogy to the activation domains of transcription factors would support the latter scenario. An approach used in the analysis of transcription activation domains was to generate a library of fragments (derived from *E.coli* chromosomal DNA) and to show that in yeast various different acidic peptides possessed transactivation activity when fused to a DNA-binding domain (Jun and Ptashne, 1987). A random rather than directed approach could be adapted to *PRP8*, utilising pAJY22, to determine what other peptide sequences possess N-terminal domain function.

It may be possible, using with some of the strains and constructs described in this thesis, to design a genetic screen to isolate the factor or factors which interact with the Prp8p N-terminal domain. For example, the defects caused by truncation of the N-terminal domain may be suppressed from high-copy libraries, or by mutagenesis, or alternatively a synthetic lethal interaction with a truncation mutant may be sought.

A commonly-used procedure for identifying factors which bind to a particular domain is to screen a fusion library by the 2-hybrid method (Chien *et al.*, 1991). This procedure is usually carried out with a library fused to an activation domain, and the gene-of-interest (or domain) fused to the DNA binding domain (as the "bait"). A requirement is thus that the polypeptide fused to the DNA binding domain does not have intrinsic transcription activating activity. Given the similarity both in sequence and in function between the N-terminal domain and known activation domains of transcription factors, it is probable that this particular screening approach will be

precluded. An alternative might be to use a systematic approach, and test known splicing factors in turn for two-hybrid interaction with this domain.

5.7.4 The Role in Splicing of the Prp8p N-terminal Domain.

The N-terminal Domain of Prp8p is assigned a general role in promoting assembly processes involving Prp8p and the U5 snRNP. From the results presented in this thesis, it is not possible to describe the mechanism by which this takes place. However, the behaviour of the N-terminal domain is typical of a proline-rich region, and as information is revealed about the modes of action of other proline-rich domains such as activation domains of transcription factors, the mechanism of snRNP assembly and the contribution of this region of Prp8p may become more clear.

There are several possible reasons why yeast Prp8p may possess an extra proline-rich domain not common to metazoans. Perhaps the low concentrations of splicing factors in the yeast nucleus imposes a requirement for extra activities promoting association and assembly, one of which is the Prp8p N-terminal domain. Alternatively, the role of the Prp8p N-terminal domain in yeast may be played by another splicing factor in metazoans. It is interesting to note that SAP62, the mammalian homologue of Prp11p has an extra proline-rich (C-terminal) domain not common to Prp11p (Bennet and Reed, 1993). The SAP62 C-terminal domain is much longer than the Prp8p N-terminal domain, with 22 heptad repeats, but remarkably contains the same LP_nG motif (a single copy of LP_5G) and also the sequence LPPPAPG adjacent to the heptad repeats. We speculate that Prp8p and Prp11p may interact with each other, or a third party, via their proline-rich domains.

It has been suggested that a function of the RNA polymerase II large subunit C-terminal domain is to interact non-stoichiometrically with other components of the transcription initiation complex to limit their diffusion away from region of transcriptional activity (Sigler, 1988). The analogy to Prp8p would propose a role for the N-terminal domain in maintaining the U5 snRNP (or the U4/U6.U5 tri-snRNP particle) in an appropriate position. For example, the domain may maintain an association with the U2 snRNP when the snRNPs are not assembled into an active spliceosome.

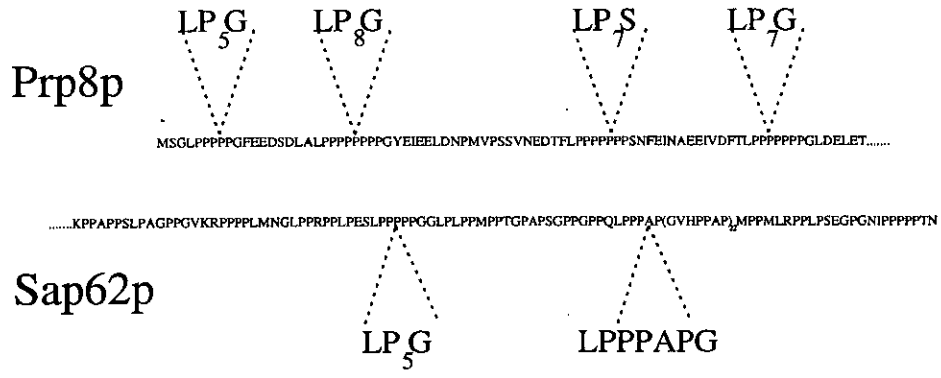


Figure 5.5: LP_nG and similar motifs in the N-terminal domain of yeast Prp8p and the C-terminal domain of mammalian SAP62, the Prp11p homologue.

Chapter 6.

The *dbf3-1* and *dna39* Alleles of *PRP8*.

6.1: Introduction.

The *dbf* (dumbbell former) yeast mutants were isolated from a collection of randomly generated temperature-sensitive mutants on the basis of an abnormal morphology, which suggested a defect in cell cycle progression (Johnston & Thomas, 1982a). The dumbbell morphology is caused by a defect which blocks the cell cycle after initiation of a daughter bud, but allows growth to continue. Bud-emergence is coupled to the start of S-phase, so *dbf* mutants are expected to affect cell cycle progression between the beginning of S-phase and cytokinesis. Alleles falling into four complementation groups, *dbf1* to *dbf4*, all seemed to specifically affect DNA replication in that synthesis of DNA ceased on switching to the restrictive temperature of 37°C, whereas synthesis of RNA and protein continued (Johnston & Thomas, 1982a). Furthermore, *dbf3* and *dbf4* were shown to be defective in S-phase initiation (Johnston & Thomas, 1982b). Thus the *DBF* genes were expected to have roles in cell cycle progression, and *DBF3* and *DBF4* specifically in S-phase and its control.

These expectations have proved well-founded for *DBF2* and *DBF4*. The product of *DBF4* interacts with the Cdc7 kinase *in vivo* to regulate its activity (Dowell *et al.*, 1994; Jackson *et al.*, 1993; Kitada *et al.*, 1992). Cdc7p is thought to act on the ORC complex which exists bound to DNA at origins of replication (Rao & Stillman, 1995) and which is required for the initiation of DNA synthesis (Foss *et al.*, 1995; Bell *et al.*, 1995; for review see Toyn *et al.*, 1995). *DBF2* also has a regulatory role in the cell cycle. Dbf2p is a kinase, which is required during nuclear division after the metaphase/anaphase boundary. At this time the kinase activity becomes de-repressed, and concomitantly the protein itself undergoes dephosphorylation (Toyn and Johnston, 1994).

The *dbf3-1* mutant has been subjected to extensive analysis in relation to its cell cycle arrest phenotype (Shea *et al.*, 1994). Heat-shock to the non-permissive

temperature delays the onset of S-phase by 45 minutes. After this time a cycle of DNA replication proceeds, but is apparently incomplete and does not in all cases generate two viable sets of chromosomes. Execution point analysis suggests that *DBF3* function is required around the initiation of S-phase.

The *DBF3* gene was cloned by complementation of thermosensitivity and sequenced entirely (Shea *et al.*, 1994). When the sequence of *PRP8* (Hodges *et al.*, 1995) became available in the database, the identity of the two was revealed. The published sequences of *DBF3* and *PRP8* (EMBL accession numbers L29421 for *DBF3*, and U00027 and Z24732 for *PRP8*) differ at four nucleotide positions (Hodges *et al.*, 1995) which most likely correspond to errors in the *DBF3* sequence. The identity was unexpected after the indirect evidence that *dbf3-1* did not affect splicing: *dbf3-1* lacks the Rna$^-$ phenotype caused by rapid rRNA turnover in the absence of ribosomal proteins whereas splicing mutants generally portray this phenotype, and also complementation studies had revealed that *dbf3-1/prp8-1* diploids are viable at 37°C. Three further alleles isolated in a separate screen for temperature-sensitive mutants in DNA synthesis are allelic to *DBF3/PRP8*; they are *dna39-1*, *-2*, and *-3* (Dumas *et al.*, 1982; Shea *et al.*, 1994). They have S-phase delay phenotypes similar, though not identical, to *dbf3-1*.

From these findings, it seemed possible that two classes of *prp8* allele were under study. Alleles of the first type affect splicing and do not arrest at a specific point in the cell cycle (Shea *et al.*, 1994), and are typified by *prp8-1*. Alleles of the second type (such as *dbf3-1*) affect the cell cycle; their function in pre-mRNA splicing had not been directly studied. This would imply that splicing factors functioned independently in a separate cellular process (cell cycle progression). Alternatively, *dbf3-1* might be unable to splice a particular, intron-containing cell cycle gene. Amongst splicing factor genes, dumbbell formation is not unique to alleles *PRP8*: yeast which harbour certain alleles of the *PRP3* gene (the *dbf5* alleles) also have this terminal morphology (Shea *et al.*, 1994). Furthermore, in the fission yeast *Schizosaccharomyces pombe* screens for mutants with cell cycle defects have revealed homologues of splicing genes (David Frendeway, personal communication). The elucidation of the basis for the relationship between pre-mRNA splicing and the cell cycle was therefore considered imperative.

6.2: Identifying the *dbf3-1* Mutation.

6.2.1: Mapping the *dbf3-1* Mutation.

Shea *et al.* (1994) showed that the mutation responsible for the *dbf3-1* phenotype in all likelihood lies within *PRP8* and not an adjacent gene. To confirm this, *dbf3-1* was mapped to a region of *PRP8* by an approach in which the chromosomal mutation is corrected by gene conversion from a cloned fragment of the wild-type gene. In this way *prp8-1* was mapped to the C-terminal 600 base pairs of the gene (Jackson, 1988), a region subsequently shown to contain a point mutation responsible for the phenotype (Hodges *et al.*, 1994; see also Section 7.3). The *prp8-2*, *prp8-5* and *prp8-7* mutations have also been mapped (Hodges *et al.*, 1994).

A set of four subclones in the multi-copy vector pFL44 which together span almost the entire *PRP8* gene has been generated (Figure 6.1; Brown, 1992). The procedure mapped *dbf3-1* to a region of the gene encompassed by pJDY4 but not the overlapping regions of pJDY2 or pJDY3. Of a second series of plasmids together spanning the region unique to pJDY4, pAJY18 caused an increase in reversion frequency, and pAJY19 and pAJY20 did not. This mapped *dbf3-1* to between positions 4094 and 4553, on an *Eco*RI-*Xba*I fragment of a size amenable to sequencing (Figure 6.1). As has been observed previously, the degree to which reversion was stimulated was roughly proportional to the size of the fragment borne on the plasmid (Jackson, 1988). Presumably, this was because larger fragments constitute more extensive substrates upon which the mechanisms of homologous recombination may act.

6.2.2: Sequencing the *dbf3-1* Mutation.

In order to identify the mutation within the *Eco*RI-*Xba*I region of the *dbf3-1* allele, a fragment was amplified by the polymerase chain reaction, using genomic DNA extracted from haploid strain L149-7B (*dbf3-1*) as the template. Oligodeoxynucleotide primers M5583 and M5584 were designed to yield the *Eco*RI-*Xba*I region plus about 30 nucleotides on either side. The amplified fragment was digested with *Eco*RI and *Xba*I, and cloned into these sites in the polylinker of pTZ18R. Single-stranded DNA was sequenced using 819X and M13 reverse oligodeoxynucleotide primers. The insert

Figure 6.1: Mapping the *dbf3-1* Mutation.

To assign the mutation to a region of *PRP8*, plasmids were introduced into L149-7B (*dbf3-1*), and transformants grown twice to stationary phase (23°C, in YMGcas supplemented with tryptophan), maintaining selection for the plasmid. Cultures were then spread to YPDA plates at various dilutions and incubated at 23°C or 37°C. The frequency of reversion was calculated as the number of revertants growing at 37°C as a proportion of viable cells. Aggregate results from at least two experiments in each case are presented in the table below. Plasmids carrying fragments spanning the position of *dbf3-1* stimulated reversion compared to controls, due to recombination allowing repair of the mutant allele.

Plasmid	Region of *PRP8* (nucleotide position)	Reversion frequency
pFL44	[vector]	$<1x10^{-7}$
pJDY1	1 to 1715	$<1.5x10^{-7}$
pJDY2	1715 to 4097	$<1x10^{-7}$
pJDY3	5329 to 7225	$<1x10^{-7}$
pJDY4	3363 to 5764	$3x10^{-5}$
pAJY18	4097 to 4554	$3x10^{-6}$
pAJY19	4554 to 5082	$<1.25x10^{-7}$
pAJY20	5082 to 5329	$<1x10^{-7}$

Key *PRP8* coding sequence is indicated by the thick line, with scale bars above and restriction sites below. L: *Sal*I, B: *Bgl*II, R: *Eco*RI, X: *Xba*I, C: *Cla*I. ✓ indicates a stimulation of reversion frequency; X indicates no stimulation.

contained a single transition mutation, C to T on the coding strand, and no other deviation from wild-type *PRP8*. To confirm that this mutation was not a PCR error, and indeed derived from the L149-7B (*dbf3-1*) genome, single-stranded DNA from independent isolates deriving from separate PCR reactions was also sequenced with 819X primer and shown to contain the same C to T mutation. Thus *dbf3-1* carries a missense mutation affecting amino acid 1489 (Figure 6.2), changing proline (CCT) for leucine (CTT).

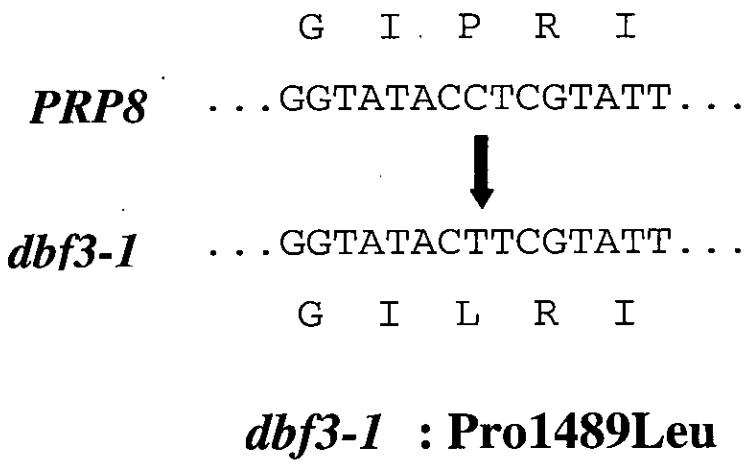


Figure 6.2: Nucleotide sequence of wild-type *PRP8* and the *dbf3-1* allele, and encoded peptide sequence. The position of the *dbf3-1* point mutation is indicated by arrow.

6.3: Effects of the *dbf3/dna39* Alleles on Splicing.

Indirect evidence that *dbf3-1* does not affect splicing was obtained in the original screen of mutants, in which *dbf3-1* was shown not to reduce net RNA synthesis (Johnston and Thomas, 1982a). Splicing defects can be examined more directly and more readily through their effects on *in vivo* levels of splicing substrates and products, and on *in vitro* splicing assays. By these means *dbf3-1* and the *dna39* alleles were analysed to determine directly if they affect splicing.

6.3.1 *In vivo* Splicing Defects of the *dbf3/dna39* Mutants.

Figure 6.3A presents a Northern analysis of the splicing phenotype of *dbf3/dna39* cells. Total RNA was extracted from cells grown at either the permissive temperature (23°C) or after heat-shock to the restrictive temperature (36°C for *dna39*-1 & -2 and *cdc25-1* cells, 37°C for *dbf3-1* and *dna39-3* cells) and subjected to Northern blotting. RNA prepared from a temperature-sensitive strain defective in *CDC25*, a gene not directly involved in splicing, serves as a negative control.

Cells of the *cdc25-1* strain contain similar levels of mRNA derived from the intron-containing *RP28* gene at both temperatures (Figure 6.3A, lanes 1 and 2). The modest reduction which is apparent may reflect the rapid and transient reduction of transcription of ribosomal protein genes which occurs upon heat shock (Larkin *et al.*, 1987; Schwindinger and Warner, 1987 and references therein). In contrast *dbf3/dna39* cells become substantially depleted of *RP28* mRNA upon heat shock. In the case of *dna39-1, dna39-2,* and *dna39-3,* pre-mRNA accumulation occurs concomitant to the mRNA depletion and these strains can be said to have the classical splicing defect affecting the mRNA:pre-mRNA ratio.

The situation of L149-7B (*dbf3-1*) is more complicated: *RP28* mRNA is depleted but no concomitant accumulation of pre-mRNA can be detected. This observation was reproduced with further preparations of total RNA extracted at various times after heat-shock of cultures to 37°C. At the restrictive temperature, mRNA from the intron-containing genes *ACT1* (data not presented), *RP28* and *CYH2* (Figure 6.3B) was depleted, whereas mRNA from *DED1*, a gene which lacks an intron, was unaffected. Again, no pre-mRNA was detected, and consideration of mRNA:pre-mRNA ratio is inappropriate. As depletion of mRNA is specific to intron-containing transcripts, these data suggest that *dbf3-1* also affects splicing.

With L149-7B, the phenotype of mRNA depletion in the absence of precursor accumulation is reminiscent of two other *prp8* strains, JDY8.57 (*prp8-5*) and JDY8.31 (*prp8-3*), neither of which accumulate substantial levels of precursor at the restrictive temperature. In Chapter 7, evidence is presented that in these strains and in L149-7B an aspect of the strain background which is not closely linked to the *PRP8* locus destabilises pre-mRNA (relative to other strains). When outcrossed into a different

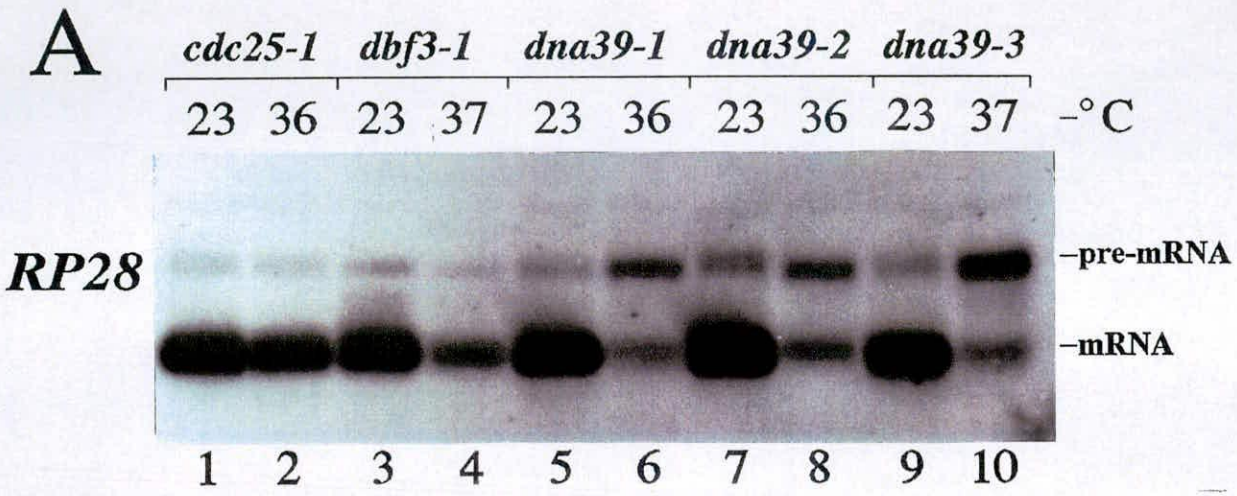
Figure 6.3: *dbf3/dna39* Alleles Affect Splicing *In vivo*.

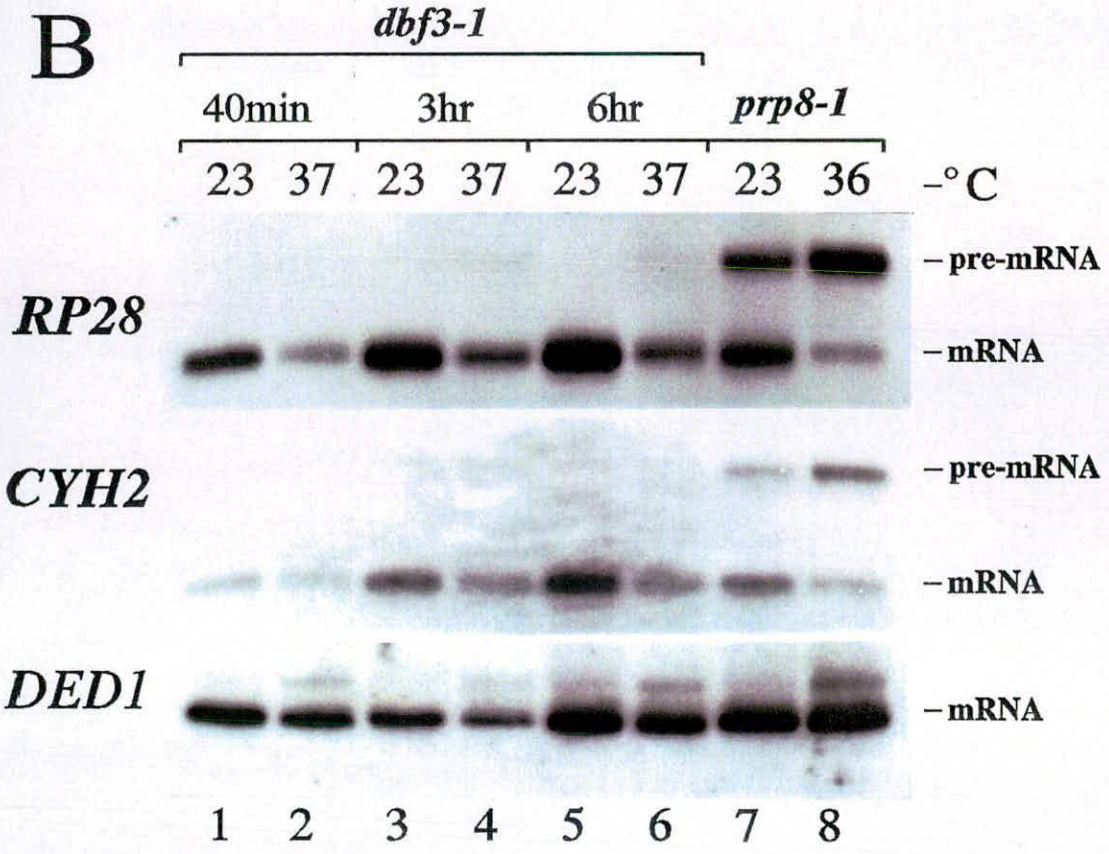
(A): The *dna39* alleles cause pre-mRNA accumulation.

Mid-log cultures in YPDA of Y166 (*cdc25-1*), JL1U (*dna39-1*), JL2U (*dna39-2*), JL3U (*dna39-3*), and L149-7B (*dbf3-1*) cells were diluted into equal volumes of fresh pre-aerated medium prewarmed to induce heat-shock to the restrictive temperature (36°C or 37°C) or maintain the permissive temperature of 23°C. Incubation was continued for 1hr at the restrictive or permissive temperature, then total RNA was prepared and analysed for RP28 transcripts as described in Figure 4.2C.

(B): *dbf3-1* causes depletion of spliced mRNA.

Mid-log cultures in YPDA of L149-7B (*dbf3-1*) cells were transferred to 37°C or maintained at 23°C (as above); incubation was continued for 40min (lanes 1 and 2), 3hr (lanes 3 and 4), or 6hr (lanes 5 and 6) and total RNA prepared. As an experimental control, total RNA was prepared from SPJ8.31 (*prp8-1*) cells treated as described in (A). Total RNA was analysed by Northern blotting as described in Figure 4.2C, hybridising the blot sequentially to 32P-labelled DNA fragments encoding *RP28*, *CYH2* and *DED1* (also *ACT1* and *LYS2*, not shown). The *RP28*, *CYH2* and *ACT1* genes contain introns (positions of pre-mRNA and mRNA shown); *DED1* and *LYS2* do not contain introns, and the signals for their mRNAs were used to control loading. The identity of the species which hybridises to the *DED1* probe and migrates slightly above *DED1* mRNA is unknown; it may represent a rare alternative transcriptional start site or polyadenylation signal at the *DED1* locus.


A
cdc25-1
dbf3-1
dna39-1
dna39-2
dna39-3
23 36 23 37 23 36 23 36 23 37 -°C
RP28
-pre-mRNA
-mRNA
1 2 3 4 5 6 7 8 9 10


B
dbf3-1
40min
3hr
6hr
prp8-1
23 37 23 37 23 37 23 36 -°C
RP28
-pre-mRNA
-mRNA
CYH2
-pre-mRNA
-mRNA
DED1
-mRNA
1 2 3 4 5 6 7 8

genetic background, *dbf3-1* is associated with the classical splicing defect of perturbation to the pre-mRNA:mRNA ratio (Figure 7.2). Thus all four *dbf3/dna39* alleles cause splicing defects *in vivo*.

6.3.2: *In vitro* Splicing Defects of the *dbf3-1* Mutant.

Whole-cell yeast extracts are able to accurately process synthetic precursor mRNA *in vitro*. The activity of extracts from *prp* mutant strains mimics the phenotype seen *in vivo* in that splicing activity is heat labile. This originally enabled the conclusive assignation of the products of *PRP* genes to the splicing process (Lustig *et al.*, 1986). To confirm that the temperature-sensitive *dbf3-1* mutation affects splicing, a similar analysis was carried out using extracts prepared from the L149-7B (*dbf3-1*) strain. In the experiments presented in this section each splicing reaction was divided into two to examine by electrophoresis both the splicesomal complexes formed (on non-denaturing gels), and the RNA intermediates and products of splicing (on denaturing gels). Results of these two analyses were in accordance in that when complex II, representing the active spliceosome, was detected lariat intermediates and spliced product also occurred. Only the non-denaturing gels are presented here.

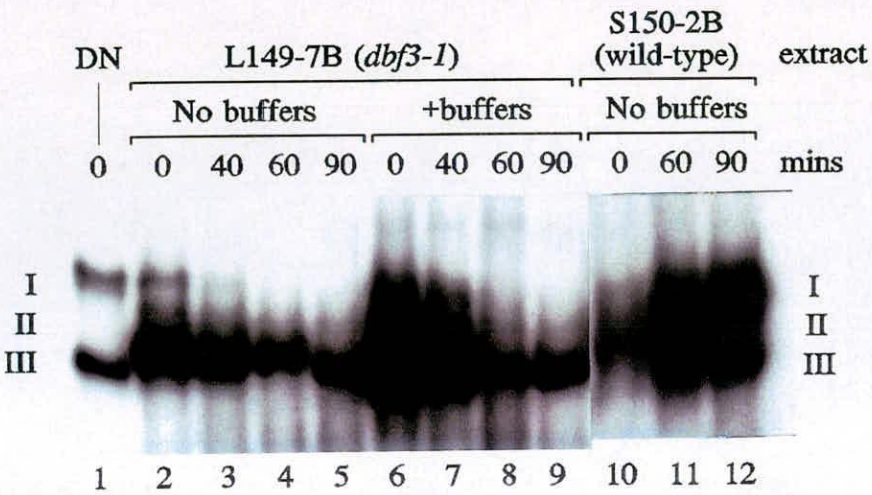
Splicing extract from the *dbf3-1* strain L149-7B was active in the *in vitro* splicing assay (Figure 6.4A lanes 2 & 6 and data not shown). Treatment of this extract at 37°C before carrying out the splicing reaction progressively abolishes activity. After 60 minutes at 37°C no complex I or II was detected (lane 4). The prespliceosome, complex III, was still detected and thus *dbf3-1* affects the same stage of spliceosome assembly as is affected by *prp8-1* (Brown and Beggs, 1992). In the control wild-type extract splicing proceeds close to its maximal extent under the conditions of the assay, as evidenced by the production of spliced mRNA (not shown) and the presence of complex II but not the complexes which precede it in the assembly pathway (I and III; Figure 6.4A lane 10). This is expected from the known kinetics of the wild-type reaction (Pikielny *et al.*, 1986). After heat-treatment of the wild-type extract complexes I and III were detected in addition to complex II (lanes 11 & 12) and splicing proceeded to a lesser extent, reflecting the stringency of inactivation conditions. However the conditions which completely abolish *dbf3-1* activity only partially inhibit the wild-type extract.

Figure 6.4: Inactivation and Complementation of *dbf3-1* Splicing Extract.

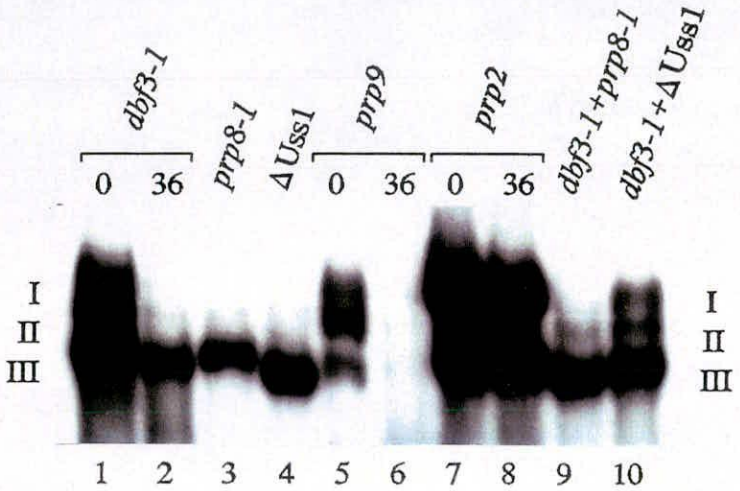
(A): Inactivation of *dbf3-1* extract. Splicing extract from L149-7B (*dbf3-1*) or wild-type strain S150-2B was pre-treated for 90min before carrying out standard splicing reactions (20µl) with radiolabelled rp28 synthetic intron-containing transcript. For the pre-treatment, extract was held on ice for 90min (lanes 2 and 10) or incubated at 37°C for 40min (lane 3), 60min (lanes 4 and 11) or 90min (lanes 5 and 12), and then placed on ice for the remainder of the 90min. For lanes 6 to 9, extract was mixed with splicing buffers (all components of the splicing reaction except transcript), and before pre-treatment. After the splicing reaction (25°C, 30min), half of each reaction was fractionated by electrophoresis through a composite 0.25% agarose, 3% polyacrylamide non-denaturing gel. Lane 1 is for reference, containing a reaction assembled with an S150-2B extract containing dominant negative $Prp2^{dn1}$ protein (Plumpton *et al.*, 1994). With this extract the gel system resolves complex III (the prespliceosome) and complex I, but not complex II, the active spliceosome (designation according to Pikielny *et al.*, 1986).

(B): Biochemical Complementation of *dbf3-1*. Extracts were prepared from L149-7B (*dbf3-1*), J17/J93 (*prp9-1*), and DJY85 (*prp2-1*) strains grown at the permissive temperature, and also from MCY4 (*GAL1:USS1*) grown in glucose to deplete Uss1 protein (Cooper *et al.*, 1994), and from SPJ8.31 heat-treated sphaeroplasts (Brown and Beggs, 1992). The active extracts (*dbf3-1*, lane 1; *prp9-1*, lane 5; *prp2-1*, lane 7) were inactivated by incubation at 36°C (1.5hr for *dbf3-1*, 1hr otherwise). Splicing reactions (20µl) were assembled with radiolabelled rp28 transcript and single extracts (10µl; lanes 1 to 8) or for biochemical complementation (lanes 9 to 18) with pairwise mixtures (5µl + 5µl) of inactive extracts. Half of each reaction was fractionated by electrophoresis through a composite 0.25% agarose, 3% polyacrylamide non-denaturing gel. Designation of the complexes I, II and III is according to Pikielny *et al.* (1986).

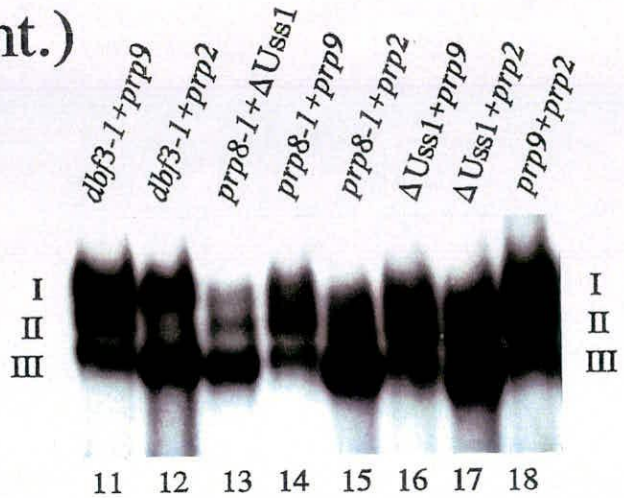
A


DN
L149-7B (dbf3-1)
S150-2B (wild-type)
extract
No buffers
+buffers
No buffers
0
0 40 60 90 0 40 60 90 0 60 90
mins
I
II
III
1 2 3 4 5 6 7 8 9 10 11 12

B


dbf3-1
prp8-1
ΔUss1
prp9
prp2
dbf3-1+prp8-1
dbf3-1+ΔUss1
0 36 0 36 0 36
I
II
III
1 2 3 4 5 6 7 8 9 10

B (cont.)


dbf3-1+prp9
dbf3-1+prp2
prp8-1+ΔUss1
prp8-1+prp9
prp8-1+prp2
ΔUss1+prp9
ΔUss1+prp2
prp9+prp2
I
II
III
11 12 13 14 15 16 17 18

To determine if less stringent conditions also inactivated *dbf3-1* extract, inactivation was carried out at various temperatures (30°C to 37°C) for 1 hour or 1.5 hours (data not presented). The mutant extract was reproducibly inactivated by treatment at 36°C. Lower temperature treatments (for example, 33°C) were effective in some but not all experiments. The addition of splicing buffers and ATP prior to heat inactivation, as was effective with the alleles studied by Lustig *et al.* (1986), did not enhance inactivation (Figure 6.4A). Thereafter, *dbf3-1* was inactivated *in vitro* by direct incubation of extracts for 1.5hr at 36°C (without premixing with splicing buffer).

Heat inactivated extracts from different *prp* mutant strains can be combined to reconstitute activity and demonstrate that the heat treatments specifically affect different exchangeable factors essential for splicing activity (Lustig *et al.*, 1986). Biochemical complementation fails when the mutations are allelic or are in genes whose products are not exchangeable (for example *PRP3* and *PRP4* deficient extracts complement poorly, Lustig *et al.*, 1986). Such an experiment with the inactive *dbf3-1* extract is presented in Figure 6.4B. The inactive *dbf3-1* extract could be complemented by extracts lacking Prp2p, Prp9p or Uss1p activity, but not by an inactive extract from SPJ8.31 (*prp8-1*). Each inactive extract was blocked at a stage of the splicing pathway consistent with the respective mutation: *prp9-1* prevented prespliceosome assembly, *prp8-1* or metabolic depletion of Uss1p prevented tri-snRNP addition, and *prp2-1* prevented the complex III to II transition. These results demonstrate that the heat treatment which abolishes activity of the *dbf3-1* extract specifically inactivates Prp8 protein, or a non-exchangeable component of the same snRNP particle. Also, *dbf3-1* is confirmed as an allele which affects splicing. The intragenic complementation of *dbf3-1* and *prp8-1* observed *in vivo* (Shea *et al.*, 1994) is not reproduced *in vitro*, instead *dbf3-1* behaves just as do the other *prp8* alleles which have been tested (*prp8-1*; Brown and Beggs, 1992; *prp8-2* and *prp8-7*; Brown, 1992).

6.4: The Effects of *dbf3-1* and *prp8-1* on U4/U6.U5 Tri-snRNP Formation.

Strain L149-7B (*dbf3-1*) is also sensitive to the sphaeroplast (*in vivo*) heat-treatment procedure which has been developed in this laboratory to yield inactive extracts (Brown and Beggs, 1992). L149-7B (*dbf3-1*) sphaeroplasts were incubated at

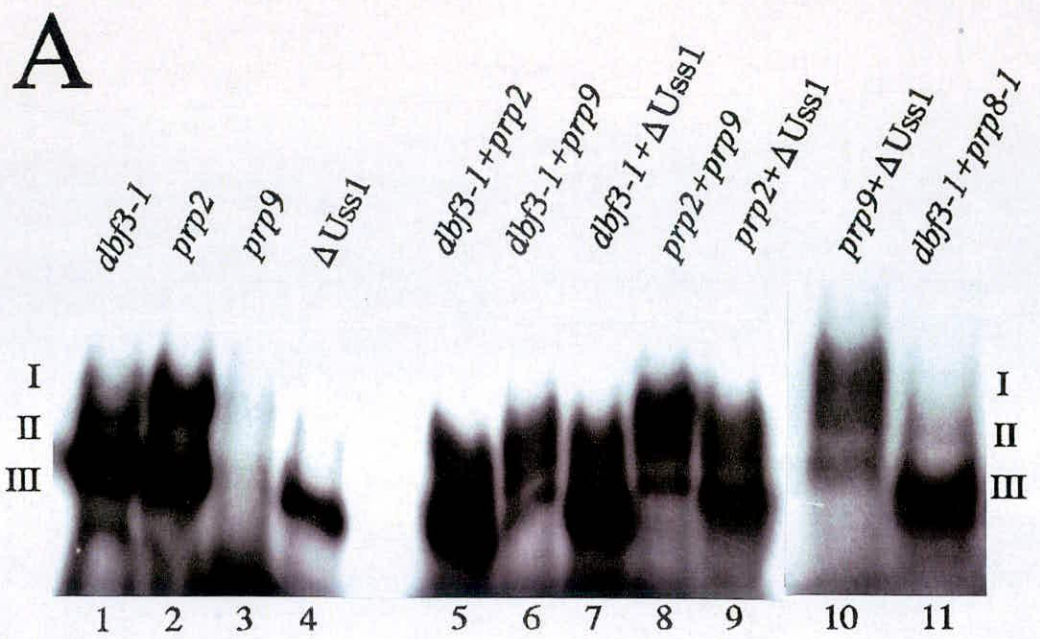
Figure 6.5: Both *dbf3-1* and *prp8-1* Influence Tri-snRNP Assembly/Stability.

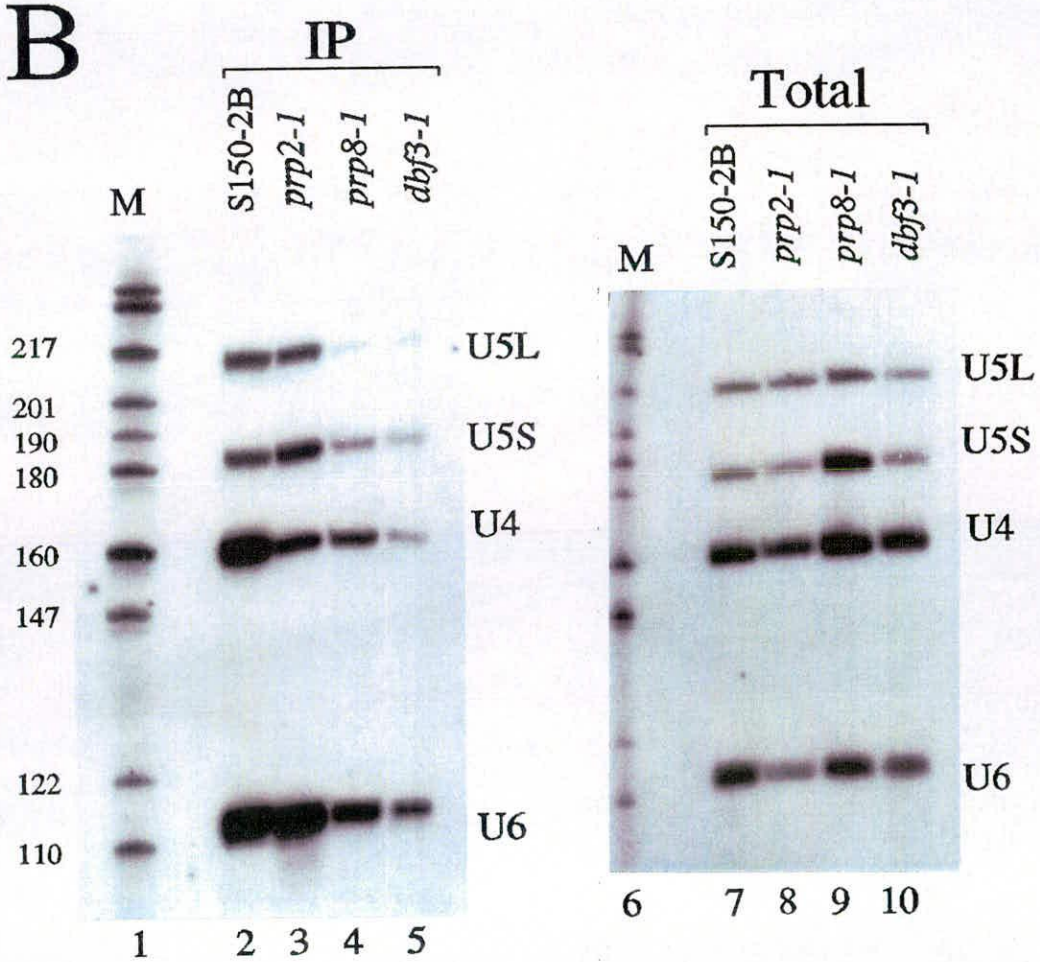
(A) Inactivation of *dbf3-1* by heat-treatment of sphaeroplasts.

Sphaerplasts were prepared from SPJ8.31 (*prp8-1*), L149-7B (*dbf3-1*) and DJY85 (prp2-1) cells and incubated for 2hr at 37°C with gentle agitation in YPDAS. Splicing extracts were then prepared according to the standard procedure. Extracts were inactive (lane 1 (*dbf3-1*), lane 2 (*prp2-1*) and lane 3 of Figure 6.4B (*prp8-1*)) but competent in biochemical complementation assays (lanes 5 to 13) performed as described for Figure 6.4B.

(B): Mutations (*dbf3-1* and *prp8-1*) reduce availability of snRNPs for coprecipitation with anti-Prp8p antibodies.

Mock splicing reactions were assembled with extracts inactivated by heat-treatment of sphaeroplasts (from A above) or with active S150-2B (wild-type) extract (1mg protein per mock reaction). SnRNP particles were precipitated with protein A Sepharose-coupled anti-8.4 immune serum, and analysed for snRNA as described in Figure 4.1. For lanes 5 to 8, complete extract (100mg) was deproteinised directly and analysed. 20% fractions of the supernate of each precipitation was also analysed (not shown) and no degradation of snRNA during the immunoprecipitation could be detected.


A
dbf3-1
prp2
prp9
ΔUss1
dbf3-1+prp2
dbf3-1+prp9
dbf3-1+ΔUss1
prp2+prp9
prp2+ΔUss1
prp9+ΔUss1
dbf3-1+prp8-1
I
II
III
1 2 3 4 5 6 7 8 9 10 11


B
IP
Total
M
S150-2B
prp2-1
prp8-1
dbf3-1
217
201
190
180
160
147
122
110
U5L
U5S
U4
U6
1 2 3 4 5
6 7 8 9 10

37°C for 2 hours in YPDAS medium; splicing extract prepared after the heat-treatment was inactive in the *in vitro* splicing assay. Biochemical complementation was possible with *prp2-1* but not *prp8-1* inactive extracts (Figure 6.5A). To determine if the *dbf3-1* mutation affected the association of Prp8p and the U4/U6.U5 tri-snRNP, immunoprecipitation was carried out using inactive extract prepared by this procedure. Mock splicing reactions were assembled with ATP but no substrate pre-mRNA, to promote maximal formation of tri-snRNP particles. Anti-8.4 antibodies precipitated U4, U5 and U6 snRNAs from the mock reactions assembled with each of the four extracts used in this experiment. However, snRNA levels were substantially reduced in the precipitates from inactive *dbf3-1* or *prp8-1* extracts, relative to an identical quantity of active S150-2B (wild-type) extract. The precipitate from the inactive DJY85 (*prp2-1*) extract contained similar quantities of the snRNAs as the wild-type control, confirming that the protoplast inactivation procedure itself does not significantly destabilise the tri-snRNP particle. Therefore the heat inactivation of *dbf3-1* which abolishes activity also destabilises the association between Prp8p and other components of the U5 snRNP particle. *dbf3-1* has very similar effects *in vitro* as *prp8-1*, in spite of their marked differences *in vivo*.

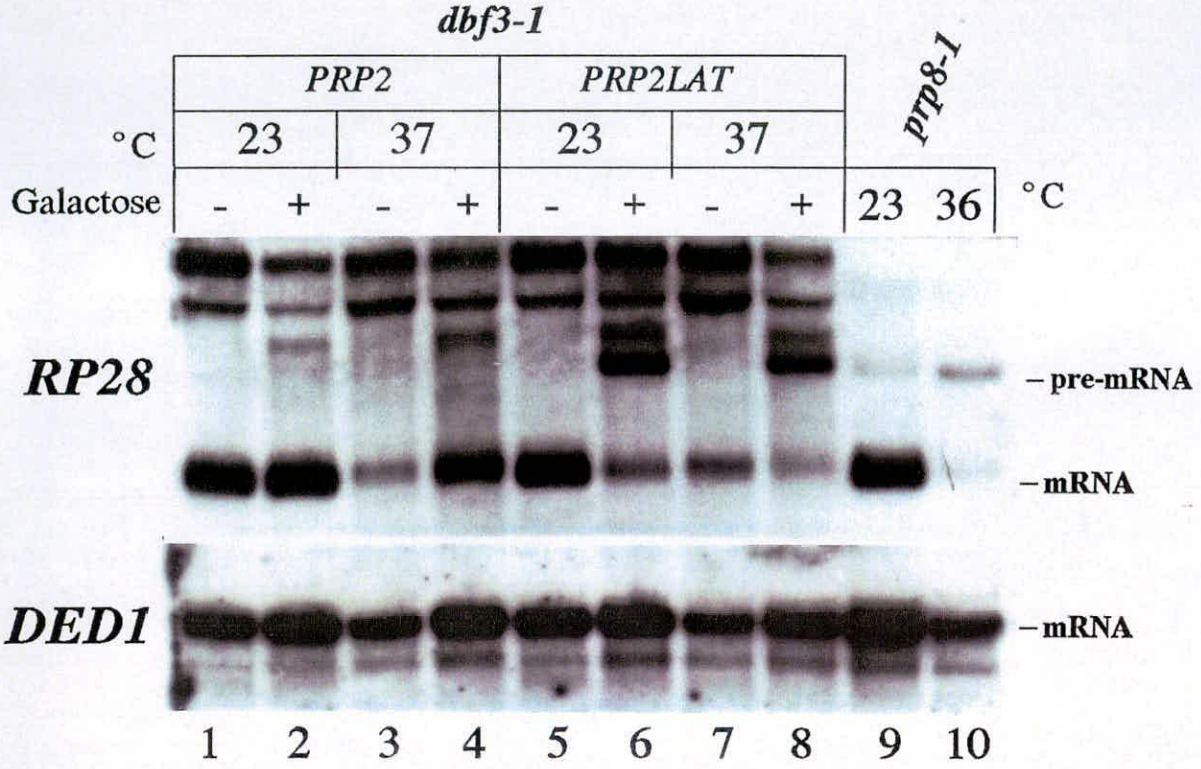
6.5: Effects of Combining of *dbf3-1* and Dominant Negative *PRP2 In vivo*.

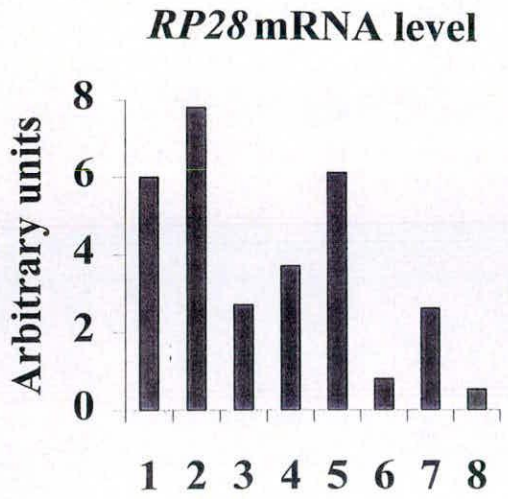
The observed splicing defects in *dbf3* and *dna39* strains are novel and unexpected. The depletion of a ribosomal protein mRNA (*RP28*) which occurs at 37°C in *dbf3-1* cells is in apparent contradiction with the previous demonstration that RNA synthesis is unaffected by this allele (Johnston and Thomas, 1982a). The two findings could be reconciled if the splicing defect of *dbf3-1* was of an intermediate nature and less severe than other *prp* mutations. According to this model, levels of ribosomal protein mRNAs in *dbf3-1* cells would be reduced to a lesser steady-state level, maintained by residual splicing activity. Ribosomal protein production would continue, and be sufficient to maintain supply of ribosomal subunits for normal rRNA processing, and rRNA turnover and net RNA synthesis would be relatively unaffected.

Some evidence in support of this model may be drawn from an experiment

Figure 6.6: *In vivo* Splicing Defects caused by *dbf3-1* and Dominant Negative *PRP2LAT* mutations.

Mid-log cultures of L149-7B (*dbf3-1*) cells bearing pBM-PRP2 (*GAL1:PRP2*; lanes 1 to 4) or PRP2^{dn1} (*GAL1:PRP2LAT*; lanes 5 to 8) in YMGlyLac (glycerol plus lactate; non-inducing) medium were mixed with equal volumes of fresh pre-warmed medium to heat-shock to 37°C (lanes 3 and 4, 7 and 8) or maintain at 23°C (lanes 1 and 2, 5 and 6). Simultaneously, galactose was added (lanes 2, 4, 6, and 8) from a sterile stock solution, or sterile distilled water (lanes 1, 3, 5, and 7). After 4hr incubation under these conditions, total RNA was prepared and analysed by Northern blotting as described in Figure 4.2C, hybridising to probes specific for *RP28* and *DED1* transcripts (presented) and *CYH2* and *ACT1* transcripts (not presented). SPJ8.31 (*prp8-1*) total RNA (lanes 9 and 10) were the same preparations as used to prepare the blot in Figure 6.3B. The species migrating slower than *RP28* pre-mRNA which hybridise to the *RP28* probe are believed to be transcripts of a nuclear *PET* gene(s) encoding mitochondrial ribosomal proteins, which are expected to be strongly induced by the glycerol/lactate carbon source. The intensity of mRNA signals were quantified by phosphorimaging. The bargraph shows *RP28* mRNA level in L149-7B (in lanes 1 to 8) after normalisation for the loading control, *DED1* mRNA.


dbf3-1
PRP2
PRP2LAT
prp8-1
°C
23
37
23
37
23
36
°C
Galactose
- + - + - + - +
RP28
– pre-mRNA
– mRNA
DED1
– mRNA
1 2 3 4 5 6 7 8 9 10


RP28 mRNA level
Arbitrary units
8
6
4
2
0
1 2 3 4 5 6 7 8

which combined with *dbf3-1* a dominant negative version of *PRP2*. This experiment was designed to analyse the nature of the pre-mRNA stability defect in L149-7B (which is not associated with the *dbf3-1* allele; Section 7.2); it also allows direct comparison of two mutations affecting splicing in the same genetic background. The *dbf3-1* allele has been shown *in vitro* to prevent assembly of the tri-snRNP particle into the forming splicesome and hence block splicing at a step prior to the involvement of Prp2p. The splicing phenotype of a complete loss-of-function *prp8* mutant is predicted not to be affected by lesion of factors such as Prp2p acting downstream in the splicing pathway, whereas the splicing phenotype of a partial loss-of-function *prp8* mutant, which allowed spliceosome assembly at some fraction of wild-type levels, might be exacerbated by a downstream lesion. The dominant negative *PRP2* allele used in this experiment was *PRP2LAT*, and was introduced into L149-7B (*dbf3-1*) on plasmid pBM-PRP2^{dn1} (Plumpton *et al.*, 1994). This plasmid contains *PRP2LAT* under transcriptional control of the *GAL1* promoter, enabling production of the toxic protein in galactose-containing growth media. *In vitro*, dominant negative Prp2^{dn1} protein competes with the wild-type for entry into spliceosomes, and when bound to the spliceosome is for some reason unable to disassociate, causing the spliceosome to stall. *In vivo*, expression of the *PRP2LAT* gene arrests growth and causes pre-mRNA accumulation (Plumpton *et al.*, 1994). The plasmid pBM-PRP2 encodes wild-type *PRP2* and was used as a control (King and Beggs, 1990). In a preliminary experiment, L149-7B (*dbf3-1*) cells bearing these plasmids were spotted to selective agar plates (YMGcas or YMGRcas, supplemented with tryptophan) and incubated at 23°C or 37°C. As expected overexpression of *PRP2* does not suppress *dbf3-1*, but overexpression of *PRP2LAT* prevents growth at 23°C of this strain (not presented).

The effects of each mutation on splicing *in vivo* was assessed by Northern analysis. To facilitate induction, cultures of the transformants were grown in glycerol/lactate (non-inducing) medium which neither induces nor represses the *GAL1* promoter. The medium lacked uracil, to maintain selection for the plasmid. Cultures were transferred to 37°C by dilution into an equal volume of pre-warmed, pre-aerated medium. The dominant negative *PRP2LAT* defect was induced by addition of galactose from a stock solution (final concentration, 2%). Four hours after transfer to 37°C and/or *PRP2LAT* induction, total RNA was extracted and fractionated by electrophoresis alongside total RNA from SPJ8.31 (*prp8-1*), to facilitate identification of mRNA and pre-mRNA bands. Transcripts of the *RP28* and *DED1* genes were identified by Northern blotting (Figure 6.6A).

Hybridisation to the probe fragment encoding *RP28* revealed multiple species in addition to *RP28* pre-mRNA and mRNA which were not present in controls (SPJ8.31 (*prp8-1*) total RNA, lanes 9 and 10). The glycerol/lactate medium used in this experiment induces aerobic, glucose derepressed growth. In contrast the total RNA of lanes 9 and 10, which does not contain the slower-migrating species, was prepared from cells grown in YPDA (ie glucose-repressed growth). The slower-migrating species were moderately abundant, and as they had sufficient homology to *RP28* to hybridise to the probe, they were possibly mitochondrial ribosomal protein-encoding mRNAs (from nuclear *PET* genes; Tzagoloff and Myers, 1986). That these transcripts were absent from cells grown under conditions of glucose-repression (YPDA, lanes 9 and 10) is consistent, as glucose derepression causes proliferation of mitochondria from dormancy (Pon and Schatz, 1991). When the same blot was hybridised to radiolabelled *ACT1* and *CYH2* DNA fragments, only mRNA and pre-mRNA bands were revealed (data not shown).

The intensity of bands on the Northern blot was quantified by phosphorimaging; *RP28* mRNA level in total RNA from L149-7B cells normalised for *DED1* mRNA is presented (Figure 6.6). The *DED1* gene is expressed constitutively at a moderate level (Struhl, 1985), and comparison to the *DED1* transcript has been used routinely as a standard for RNA recovery (Klein and Struhl, 1994). The presence of the *PRP2LAT* gene under non-inducing conditions had no detectable effect on splicing (compare lanes 1 and 5, and 3 and 7). This was expected, as active splicing extracts can

be prepared from cells bearing *GAL1:PRP2LAT* grown under the same, non-inducing conditions (Flinn, 1994). Heat-shock of *dbf3-1* reduced spliced mRNA (*RP28*) levels by approximately 50% relative to the level of an unspliced mRNA (*DED1*). The extent of this reduction is reproduced in three pairs of conditions, and is not affected by the presence of galactose (compare lanes 1 and 3, 2 and 4, 5 and 7). For comparison, heat-shock of SPJ8.31 (*prp8-1*) cells reduces *RP28* level by 4-fold (lanes 9 and 10), although in this case cells were grown in a different medium in which RNA stability is likely to differ. Overexpression of *PRP2LAT* in L149-7B cells reduced spliced mRNA levels by approximately 10-fold (lanes 5 & 6). *RP28* mRNA level resulting from heat inactivation of *dbf3-1* is further reduced by induction of *PRP2LAT* (compare lanes 7 and 8). The dominant negative *PRP2* mutant causes pre-mRNA accumulation whereas heat inactivation of *dbf3-1* does not (lanes 6 and 8); accumulation of *ACT1* and *CYH2* pre-mRNA was detected in these lanes also (not presented). These observations are consistent with a residual level of spliceosome assembly and splicing occurring in the *dbf3-1* strain, enabling the *PRP2LAT* defect to be manifested.

The impression that overproduction of wild-type Prp2p in L149-7B cells stimulates splicing, at either the permissive (Figure 6.6, lanes 1 and 2) or restrictive (lanes 3 and 4) temperature for *dbf3-1*, may be misleading. It was not expected, as wild-type levels of Prp2p are not thought to be limiting for splicing. This effect may simply be due to the presence of galactose in the medium. Increased *RP28* spliced mRNA levels (as a proportion of total RNA) upon overexpression of *PRP2* have been observed before (Plumpton *et al.*, 1994), although comparison to non-spliced standards has previously not been carried out as a matter of course.

6.6: Discussion.

The demonstration that the *dbf3-1* mutation is located within the *PRP8* coding sequence confirms the results of Shea *et al.* (1994) to exclude the possibility that *dbf3-1* is due to a separate genetic element closely linked to *PRP8*. The *dbf3-1* mutation lies in the same region of the gene as *prp8-7*. The latter allele has not yet been sequenced, but an examination of *prp8-7* cells incubated at the restrictive temperature revealed no dumbbell-like terminal morphology, so the two alleles probably contain different mutations. The amino acid change caused by *dbf3-1* is a substitution of proline by leucine, at position 1489. Proline at this position is conserved in *C.elegans* and maize. Position 1489 (yeast numbering) is in a block of nine amino acids (...SWDRGIPRI...) identical in yeast, *C.elegans* and maize, and within a region of 40 amino acids that are 75% identical and 100% similar over the three species. This high degree of conservation is not unusual within *PRP8* (other regions of the gene have even higher levels of homology), and homology alone does not delineate a domain whose function might be affected by *dbf3-1*. It may be speculated that the proline at position 1489 makes an important contribution to secondary structure in this region (perhaps as a turn or to create a kink within an α-helix) and that the mutation perturbs this structure. Secondary structure prediction (Jackson, 1988) suggested that the region preceding position 1489 (positions 1476 to 1488) had low propensity for α-helix or β-sheet formation, but predicted the five amino acids following position 1489 (1490-1495) to form a β-sheet. In proteins there is a strong tendency to find proline at the beginning of an α-helix (Richardson and Richardson, 1988), possibly because proline does not require a hydrogen-bonding partner (Williamson, 1994). Thus proline 1489 may form the N-terminal boundary of an α-helix or β-sheet structure. In the future, as more becomes known about domain structure and function within Prp8p, the molecular basis of the *dbf3-1* phenotype may become more obvious.

As expected for mutants of *PRP8*, strains bearing *dbf3-1* or *dna39* alleles are defective for splicing *in vivo* and splicing extract prepared from *dbf3-1* cells is sensitive to heat-inactivation. As was found previously with extracts from *prp8-1* cells, *dbf3-1* extracts required prolonged incubation at temperatures above 30°C to achieve inactivation. Also in common with *prp8-1* (Brown and Beggs, 1992), no set of

conditions completely and reproducibly inactivated *dbf3-1* but did not affect a wild-type extract. In general, extracts from *prp8* strains seem to be refractive to heat-inactivation and require more stringent inactivating conditions than extracts from other *prp* strains. For comparison, *prp2* extract can be inactivated by 15min incubation at 30°C (Lustig *et al.*, 1986). Biochemical complementation demonstrated that inactivation of *dbf3-1* was not non-specific, and the failure of complementation with the pair of *prp8-1* and *dbf3-1* inactive extracts is consistent with both lacking active Prp8p. In the experiment presented in Figure 6.4B, complementation between *dbf3-1* and *prp2-1* extracts was weak, although some restoration of splicing was detected from denaturing electrophoresis. In the following experiment (Figure 6.5), which used separate extracts, biochemical complementation between *dbf3-1* and *prp2-1* inactive extracts was extensive, and thus a deficiency in Prp2p activity is not a reproducible feature of inactive *dbf3-1* extracts.

In immunoprecipitation, decreased levels of U5 snRNP and U4/U6.U5 tri-snRNP were recognised by anti-Prp8p antibodies as a result of the *dbf3-1* mutation. This effect has been observed previously with the *prp8-1* mutation (Brown and Beggs, 1992). In *prp8-1* inactive extract, Prp8p appears unable to stably associate with the U5 snRNP, from its sedimentation pattern on glycerol gradients (Brown and Beggs, 1992), and the same may be true of the product of *dbf3-1* allele. Unstable association between Prp8 protein and other snRNP components has now been demonstrated for all four alleles which have been tested: *prp8-1* (Brown and Beggs, 1992), *prp8-2* and *prp8-7* (Brown, 1992), and *dbf3-1* (this study). The results of this experiment differ from those obtained previously in that before no snRNP particles were detected to be co-precipitated from inactive *prp8-1* extract, whereas in this study Prp8p-containing snRNPs could still be detected. In this study immunoprecipitation was carried out on mock splicing reactions, whereas previously immunoprecipitation was on the extract incubated in the absence of ATP and splicing buffers. This, or variability between extracts, could account for the difference. The data presented here imply that *prp8-1* inactivation and Prp8p release from (or failure to associate with) the U5 snRNP are not obligatorily associated, and that snRNP particles occur which contain Prp8p, but which are inactive due to *prp8-1*.

From the immunoprecipitation data, the failure of *dbf3-1* and *prp8-1* inactive extracts to complement when mixed is apparent: both are deficient in the association of Prp8p with the U5 snRNP. An alternative approach in attempting to reproduce *prp8-1/dbf3-1* intragenic complementation *in vitro* would be to mix and preincubate active extracts, and then treat at higher temperatures, with as a control each extract in isolation treated likewise. Of course the difficulty with this is that *prp8-1* and *dbf3-1* extracts are relatively refractive to inactivation. This has not yet been attempted.

In conclusion, from *in vitro* assays the phenotypes of *dbf3-1* and *prp8-1* are quite similar: both mutations affect the stage of tri-snRNP assembly and the association between Prp8p and components of the U5 snRNP. This is remarkable because behaviour of the two alleles *in vivo* differs: *prp8-1* results in an Rna$^-$ phenotype and growth arrest at no particular stage of the cell cycle, whereas *dbf3-1* is not associated with a Rna$^-$ phenotype but results in growth arrest at a specific cell cycle stage (S-phase). The contrast between *in vitro* and *in vivo* results is not paradoxical if the *dbf3-1* splicing defect is of mild severity and less severe than *prp8-1*. The fact that *dbf3-1* portrays a cell cycle phenotype at all implies a mild severity splicing defect, because a mutation which resulted in complete inactivation of splicing would result in coordinate depletion of actin, ribosomal proteins, and numerous other essential cellular components, and presumably quite rapid cessation of growth. The arrest of *prp8-1* cells at random stages of the cell cycle (Shea *et al.*, 1994) is in keeping with this. *dbf3-1* cells transferred to the restrictive temperature just after S-phase must be able to proceed through G2, M, and G1 phases and continue to grow, until arresting in the next S-phase. The Rna$^+$ phenotype of *dbf3-1* strains is also evidence of a mild severity splicing defect as remarked previously.

The sensitivity of *dbf3-1* to inactivation *in vitro* does not contravene this model, because the *in vitro* inactivation process is optimised to yield inactive extract. Thus the process of inactivation may only occur to a limited extent *in vivo* but occur to completion when *dbf3-1* extracts are heat-treated *in vitro*, under the optimised condition. The inactivation process also occurs in sphaeroplasts, where presumably the splicing apparatus still resides in an intact nucleus. However, to what extent the sphaeroplast heat-treatment mimics transfer of logarithmically growing cells to the restrictive temperature is not known. Certainly, the metabolic state of the sphaeroplast

is different. If the *dbf3-1* splicing defect is indeed of mild severity *in vivo* this would imply that the sphaeroplast treatment is not representative of heat-treatment during normal growth. It may in future be informative to assess the extent of inactivation of Prp8p in splicing extracts prepared from *dbf3-1* cells incubated at the restrictive temperature (and *prp8-1* for comparison).

Direct quantitation of Northern blots suggests that transfer of *dbf3-1* cells (grown in YPDA) to the restrictive temperature reduces *RP28* spliced mRNA level by 2 to 3-fold, whereas other *prp8* mutations including *prp8-1* and *prp8-7* reduce *RP28* spliced mRNA by 4-fold or more. However, these are steady-state levels, influenced by rates of mRNA degradation as well as rates of splicing, and potentially sensitive to variation due to different genetic backgrounds. As an isogenic pair of strains bearing *prp8-1* and *dbf3-1* is not available, ectopic induction of a second (dominant) splicing defect was used to compare different splicing defects in the same genetic background. The dominant mutant utilised was the SAT to LAT dominant negative mutant of *PRP2*. A limitation of this experiment is that the two mutations were induced by different means, *PRP2LAT* by transcriptional upregulation, and *dbf3-1* by heat-shock. Feasibly, the defects may have different time courses with rapid inactivation of Prp8p but gradual accumulation of toxic Prp2p (or vice versa). The interpretation that *dbf3-1* must allow residual splicing to enable *PRP2LAT* to have any additional effect relies on the extrapolation of behaviour *in vitro* to the situation *in vivo*. Also, a similar experiment in a different *prp8* strain for comparison has not been carried out. However, taken in combination with the other evidence, the model of a mild severity splicing defect of *dbf3-1* is quite well supported.

Two alternative models to explain the cell cycle phenotype of *dbf3-1* were introduced in Section 6.1. Prp8p may possess two sets of functions, one in pre-mRNA splicing and one concerned with cell cycle progression, with the two types of function differentially affected by different *prp8* mutations. Alternatively, Prp8p may directly function only in splicing such that *PRP8* would be dispensable if yeast lacked introns. In the latter scenario, different phenotypic effects of *prp8* mutations would be caused by differential effects on splicing. Previously, *dbf3-1* was suspected not to affect splicing due to its Rna^+ phenotype. However, the demonstration here of splicing defects associated with *dbf3-1* lends strong support to the latter hypothesis.

Assuming Prp8p only functions in splicing, then presumably the growth defect of *dbf3-1* phenotype results from a deficiency in a factor which is a) the product of splicing, and b) essential for cell cycle progression. An attempt to determine what that factor might be follows, in Chapter 8.

Chapter 7.

Intragenic Complementation Amongst *prp8* Alleles.

7.1: Introduction

This chapter describes two genetic phenomena which have been identified during these studies on *PRP8*. They have been revealed in the analysis of the set of ten temperature sensitive *prp8* alleles. Six alleles (*prp8-1, prp8-2, prp8-3, prp8-5, prp8-7* and *prp8-8*) have been outcrossed from mutagenised backgrounds in this laboratory; they derive from Hartwell's original screen for *rna* mutants (*prp8-1* to *prp8-7*) and a directed screen for mutants causing with pre-mRNA accumulation (*prp8-8*; Morran, 1990). The four *dbf3/dna39* alleles were isolated in screens for mutants affecting the cell cycle (Shea *et al.*, 1994). All ten alleles cause depletion of spliced mRNA, indicative of a splicing defect. In spite of this, three strains bearing *prp8* alleles did not accumulate the substrate of splicing, pre-mRNA, at the restrictive temperature. The evidence presented in Section 7.2 suggests that this effect is due to the strain background and not the *prp8* allele in each case, and is attributed to the wild-type strains into which the alleles were outcrossed. The mechanism by which the accumulation of pre-mRNA is prevented is unknown, but seems to have a potentiating effect in combination with truncation mutation of *PRP8*, and may have similar effects on other splicing mutations.

Section 7.3 concerns the intragenic complementation between *prp8-1* and *dbf3-1*. As has been reported (Shea *et al.*, 1994), diploids formed by mating *prp8-1* and *dbf3-1* haploid strains are not temperature-sensitive. This section describes attempts to reproduce complementation in a *dbf3-1* haploid strain with *prp8-1* plasmids which have been constructed in this laboratory.

7.2: A Precursor mRNA Instability Phenotype.

The results of Shea *et al.* (1994) were most readily explained by the inference that *prp8* alleles affected either splicing or the cell cycle. This prompted the

examination of all available *prp8* mutants for effects on splicing. The *dna39* alleles cause pre-mRNA accumulation, relative to mRNA (Figure 6.3). The *dbf3-1* allele also affects splicing, although in strain L149-7B no pre-mRNA accumulation accompanying mRNA depletion could be detected. Northern analysis of the remaining *prp8* alleles is described here.

7.2.1: Effects of *prp8* Alleles on Splicing.

As expected, incubation of strain SPJ8.31 (*prp8-1*) at the restrictive temperature (36°C) caused pre-mRNA accumulation and concomitant depletion of spliced mRNA. This phenotype is characteristic of splicing mutants, and has been demonstrated before with the same strain (Brown and Beggs, 1992). Figure 7.1 (A and B) presents Northern analysis of strains JDY8.22 (*prp8-2*), DJY87 (*prp8-7*) and JDY8.81 (*prp8-8*), and shows that these strains have the same characteristic phenotype in common with SPJ8.31 (*prp8-1*).

With the strains JDY8.57 (*prp8-5*) and JDY8.31(I1) (*prp8-3*), little pre-mRNA could be detected, although levels of spliced mRNA were depleted by incubation at the restrictive temperature, relative to loading controls (Figure 7.1A). In this respect, these strains are similar to L149-7B (*dbf3-1*; see Figure 6.3). Thus three out of ten *prp8* strains tested apparently had a combined phenotype, of pre-mRNA instability as well as splicing. The effect on pre-mRNA stability could feasibly have been caused by the *prp8* mutation. Importantly, however, pre-mRNA instability does not correlate with defective cell cycle progression. The *dna39-3* and *dbf3-1* strains both cause defective cell cycle progression (Shea *et al.*, 1994), but differ with respect to pre-mRNA stability. The *prp8-5* and *prp8-3* strains have not yet been subjected to detailed analysis in regard of cell cycle arrest, but incubation of JDY8.31(I1) and JDY8.57 cells at restrictive temperatures does not result in the formation of dumbbell-like morphologies.

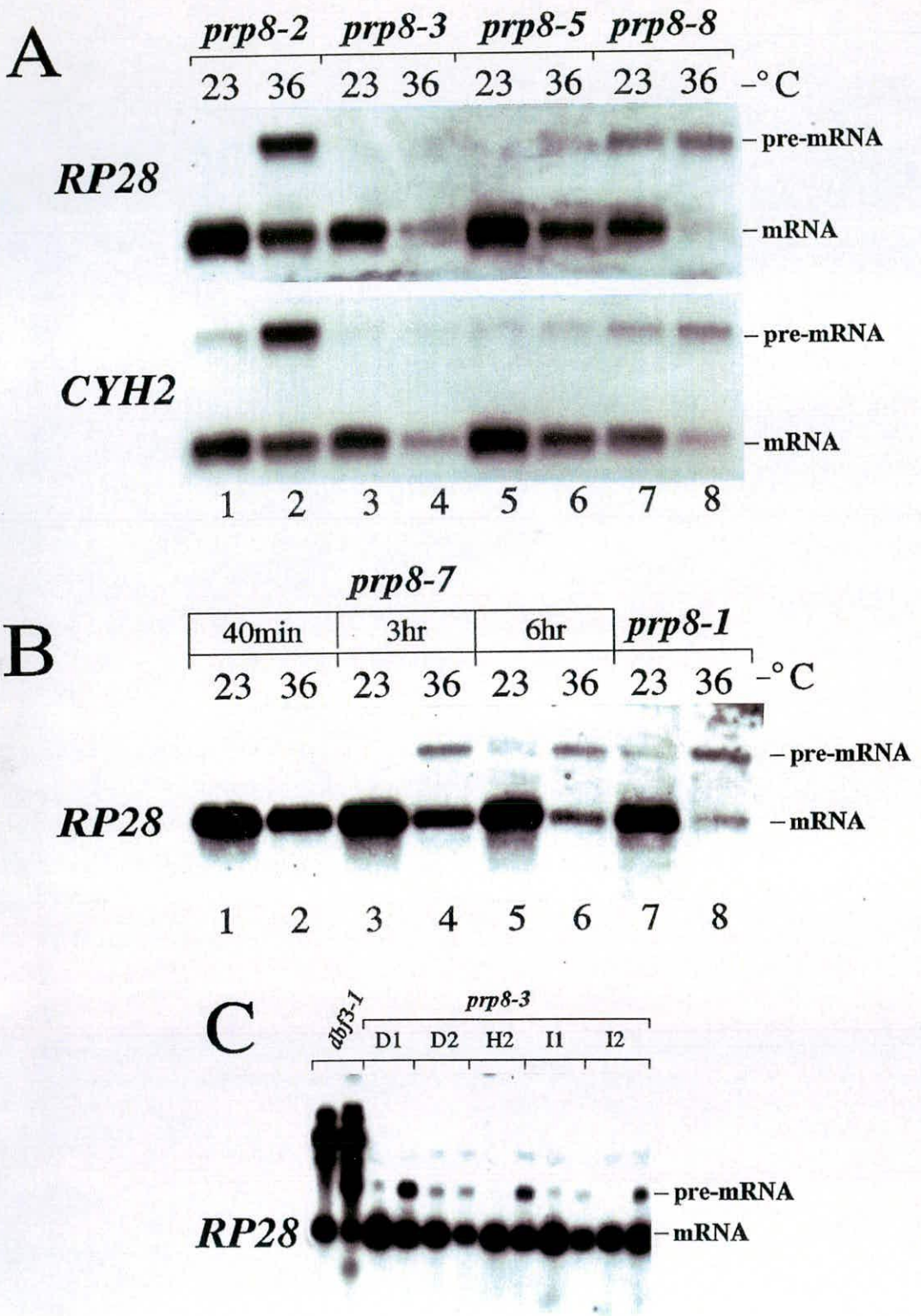
The *prp8-3* and *prp8-5* alleles originate from the Hartwell collection (Hartwell, 1967 and 1970). These and other *prp8* mutants have been outcrossed from the mutagenised background, to generate strains free of mutations other than the *prp8* allele. The *prp8-5* allele was outcrossed twice into wild-type strain DBY745, and the

Figure 7.1: *In vivo* Splicing defects of *prp8* Alleles.

(A): Pre-mRNA accumulation is not detected in *prp8-3* or *prp8-5* strains.
Total RNA was extracted from strains JDY8.22 (*prp8-2*; lanes 1 and 2), JDY8.31(I1) (*prp8-3*; lanes 3 and 4), JDY8.57 (*prp8-5*; lanes 5 and 6) and JDY8.81 (*prp8-8*; lanes 7 and 8) four hours after dilution with heat-shock to restrictive temperature (36°C; lanes 2,4,6,8) or dilution without altering temperature (23°C; lanes 1,3,5,7). RNA was analysed by Northern blotting (60µg per lane), hybridising the blot sequentially to ^{32}P-labelled DNA fragments encoding the *RP28* and *CYH2* genes. To determine variation of mRNA recovery, the blot was probed for *DED1* (not shown); yield (from 36°C cultures as a proportion of 23°C cultures) was 110% (lane 2), 40% (lane 4), 78% (lane 6), and 80% (lane 8).

(B): Pre-mRNA accumulates in DJY76 (*prp8-7*).
Total RNA from DJY76 was extracted after various times of incubation at the restrictive temperature (36°; lanes 1 to 6) as described in Figure 6.3. Northern blots were probed for *RP28* transcripts. RNA from SPJ8.31 in lanes 7 an 8 is the same as used in Figure 6.3B.

(C): Variation in pre-mRNA accumulation amongst related *prp8-3* strains.
JDY8.31(D1) (lanes 3 and 4), JDY8.31(D2) (lanes 5 and 6), JDY8.31(H2) (lanes 7 and 8), JDY8.31(I1) (lanes 9 and 10) and JDY8.31(I2) (lanes 11 and 12) are haploid *prp8-3* temperature-sensitive strains. They were generated by sporulation of diploid JDY8.31xDBY745 (*prp8-3/PRP8*) during outcrossing of the *prp8-3* allele (J.D.Brown, unpublished). Total RNA was prepared from these strains 4 hours after dilution/heat-shock of cultures, and analysed by Northern blotting for *RP28* transcripts. RNA in lanes 1 and 2 is from strain L149-7B (*dbf3-1*) bearing dominant negative *PRP2LAT* (pBM-PRP2^{dn1}) grown under inducing (lane 2) or non-inducing conditions (lane 1); it was identical to that used to prepare the blot in Figure 6.7.


A
prp8-2
prp8-3
prp8-5
prp8-8
23 36 23 36 23 36 23 36 -°C
RP28
– pre-mRNA
– mRNA
CYH2
– pre-mRNA
– mRNA
1 2 3 4 5 6 7 8
B
prp8-7
40min
3hr
6hr
prp8-1
23 36 23 36 23 36 23 36 -°C
RP28
– pre-mRNA
– mRNA
1 2 3 4 5 6 7 8
C
dbf3-1
prp8-3
D1 D2 H2 I1 I2
RP28
– pre-mRNA
– mRNA
1 2 3 4 5 6 7 8 9 10 11 12

prp8-3 allele once into DBY745 and once into S150-2B (wild-type). None of the other *prp8* alleles were outcrossed into DBY745 (Brown, 1992; D. Jamieson, unpublished). Thus the *prp8-3* and *prp8-5* strains under study have in common a parent not shared by the *prp8* strains which accumulate pre-mRNA, leading to the suspicion that the effect on pre-mRNA stability (*prp8-3* and *prp8-5* strains) derives from DBY745.

When *prp8-3* was outcrossed, other temperature-sensitive haploids from the same diploid parent as JDY8.31(I1) were isolated and stored. Figure 7.1C presents a Northern analysis of the splicing phenotype of these strains. In these strains, pre-mRNA accumulation concomitant to mRNA depletion is variable. Pre-mRNA did not accumulate in JDY8.31(I1) and in JDY8.31(D2) but significant accumulation was detected in other strains (JDY8.31(I2), JDY8.31(H2) and JDY8.31(D1)). The genotype in these strains has not been confirmed by backcrossing or otherwise, but assuming they bear the *prp8-3* allele, the implication is that the effect is not linked to the *PRP8* locus.

To confirm that pre-mRNA instability had been derived from the DBY745 background, the *prp8-1* allele was crossed into DBY745. Tetrads from the SPJ8.31xDBY745 (*prp8-1/PRP8*) diploid were dissected, and four temperature-sensitive haploid progeny deriving from different tetrads were subjected to Northern analysis (Figure 7.2B). Out of the four, one did not accumulate pre-mRNA (lanes 3 and 4). The complementary experiment was carried out by mating L149-7B (*dbf3-1*) with a wild-type strain (D150). Four spores dissected from the *dbf3-1/PRP8* diploid were designated AJY8.31, AJY8.32, AJY8.33, and AJY8.36, and subjected to Northern analysis (Figure 7.2A). AJY8.36 was temperature-sensitive at 37°C, but it was later shown that complementation by pY8000 (2μ, *PRP8*) was impossible. This, and the fact that spliced mRNA levels were relatively unaffected by transfer to 37°C (Figure 7.2A, lanes 9 and 10) suggests that its genotype is *PRP8* (wild-type) and not *dbf3-1*. AJY8.31, AJY8.32 and AJY8.33 were all temperature-sensitive at 37°C and complemented by pY8000, consistent with a genotype of *dbf3-1*. In the case of AJY8.31 (*dbf3-1*), signal intensity for both *RP28* and *CYH2* transcripts were unusually low (Figure 7.2A, lanes 3 and 4), for reasons unknown. AJY8.33 (*dbf3-1*) at 37°C was not associated with pre-mRNA accumulation, whereas AJY8.32 (*dbf3-1*) exhibited the

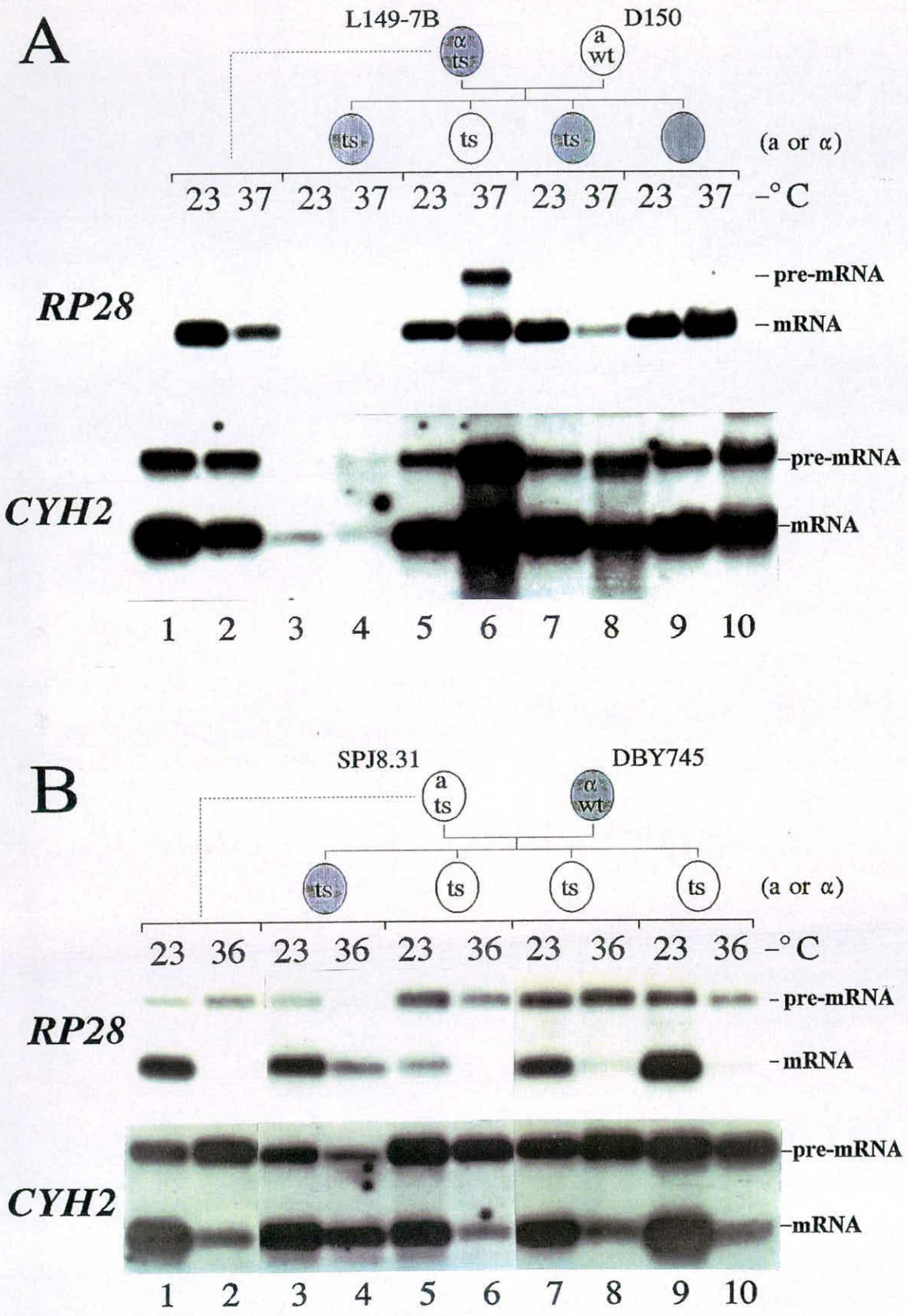
Figure 7.2: Pre-mRNA Instability is a Strain Background Effect.

(A): A *dbf3-1* strain which accumulates pre-mRNA.

AJY8.31 (lanes 3 and 4), AJY8.32 (lanes 5 and 6) and AJY8.33 (lanes 7 and 8) are haploid temperature-sensitive *dbf3-1* strains derived from L149-7BxD150 (*a/α, dbf3-1/PRP8*). Total RNA was prepared from these strains, and also L149-7B (*dbf3-1*) 4 hours after dilution into fresh medium with heat-shock to 37°C (lanes 2,4,6,8,10) or without changing temperature (23°C, lanes 1,3,5,7,9), and analysed by Northern blotting for transcripts from intron-containing genes. Strains suspected to harbour the factor which prevents pre-mRNA accumulation are shaded grey. Haploid strain AJY8.36 (lanes 9 and 10) is also derived from the L149-7BxD150 diploid, but is believed to be genotypically *PRP8* (wild-type); it is shaded, although its effect on pre-mRNA stability is not known.

(B): A *prp8-1* strain which does not accumulate pre-mRNA.

Haploid temperature-sensitive *prp8-1* strains (lanes 2 to 10) were isolated from diploid SPJ8.31xDBY745 (*a/α, prp8-1/PRP8*) and analysed along with parent SPJ8.31 (*prp8-1*) as described above.


A
L149-7B
D150
α ts
a wt
ts
ts
ts
(a or α)
23 37 23 37 23 37 23 37 23 37 -°C
RP28
-pre-mRNA
-mRNA
CYH2
-pre-mRNA
-mRNA
1 2 3 4 5 6 7 8 9 10
B
SPJ8.31
DBY745
a ts
α wt
ts
ts
ts
ts
(a or α)
23 36 23 36 23 36 23 36 23 36 -°C
RP28
-pre-mRNA
-mRNA
CYH2
-pre-mRNA
-mRNA
1 2 3 4 5 6 7 8 9 10

characteristic *prp* defect of pre-mRNA accumulation. In conclusion, the pre-mRNA stability phenotypes of *prp8-1* and *dbf3-1* strains can be swapped by crossing to alternative genetic backgrounds. Therefore, pre-mRNA instability is not caused by mutant *prp8* but by some aspect of DBY745 (and presumably M1-2B, the progenitor of L149-7B). To determine whether pre-mRNA instability is dominant or recessive, or caused by a single factor would require further spores to be subjected to Northern analysis.

7.2.2: Complementation of *prp8* Strains with Truncation Mutants.

The complementation of *prp8-1* in SPJ8.31 cells by the *prp8ΔN* truncation mutant was presented in Section 4.3. Figure 7.3 shows the complementation by truncation mutants of five other *prp8* alleles, they are: JDY8.22 (*prp8-2*), JDY8.81 (*prp8-8*), JL1U (*dna39-1*), JL2U (*dna39-2*) and JL3U (*dna39-3*). DJY76 (*prp8-7*) was also tested, and gave the same pattern (not shown). Several of these strains do not grow on galactose/raffinose carbon sources, and in these cases the plasmids which bear alleles transcribed from the *GAL1* promoter (rows 2 to 4) reveal little. However, alleles are expressed from the *PRP8* promoter irrespective of carbon source (lanes 5 to 8) and expression of *prp8ΔN* supports growth. In contrast, strains JDY8.31(I1) (*prp8-3*) and L149-7B (*dbf3-1*) were complemented by *prp8ΔP2*, but not by *prp8ΔN* (Figure 7.4A). This is not likely to be due to variable rates of transcription, because both the *GAL1* and *PRP8* promoters gave the same result. Strains bearing the same *prp8* allele as L149-7B (*dbf3-1*) differ in complementation by pAJY9 (*GAL1:prp8ΔN*): AJY8.31 and AJY8.33 (both *dbf3-1*) are not complemented by pAJY9 on galactose at the non-permissive temperature (36°C), whereas AJY8.32 is (Figure 7.4B).

It is clear that there exists a correlation amongst *prp8* strains between the stability of accumulated pre-mRNA and the ability of *prp8ΔN* to complement that strain (see Table 7.1). The hypothesis that pre-mRNA instability prevents function of *prp8ΔN* is reasonable; an alternative is that the causal relationship is not direct, but that the strains in question have mutations which both destabilise pre-mRNA and prevent complementation by *prp8ΔN*. From these results it seems likely that *prp8/prp8ΔN* intragenic complementation is exhibited by all *prp8* alleles, although for *prp8-3* and *prp8-5* this has not yet been tested in a background in which pre-mRNA is stable.

Figure 7.3: A Subset of *prp8* Alleles Complemented by *prp8ΔN*.

Yeast strains were JDY8.22 (*prp8-2*), JDY8.81 (*prp8-8*), JL1U (*dna39-1*), JL2U (*dna39-2*), JL3U (*dna39-3*). Cells bearing URA3 plasmids were transferred to plasmid-selective plates for growth at permissive (23°C) or restrictive (36°C) temperatures under inducing (YMGRcas plates, Gal) or repressing (YMGlucas plates, Glu) conditions. Plates were incubated for 5 days. Strain DJY76 (*prp8-7*) was tested in the same way (not presented); it too was complemented by pJDY7 (*prp8ΔN*). The extent of variability between strains of complementation by pBM-PRP8 under conditions of glucose repression is apparent, ranging from JDY8.31(I1) which does not grow, to JL1U, which grows strongly.

URA3 plasmids were:

(1) vector:	pBM125	(5) *PRP8*:	pJDY6
(2) *GAL1:PRP8*:	pBM-PRP8	(6) *prp8ΔN*:	pJDY7
(3) *GAL1:prp8ΔP2*:	pAJY1	(7) *PRP8*:	pAJY12
(4) *GAL1:prp8ΔN*:	pAJY9	(8) *prp8ΔP2*:	pAJY13

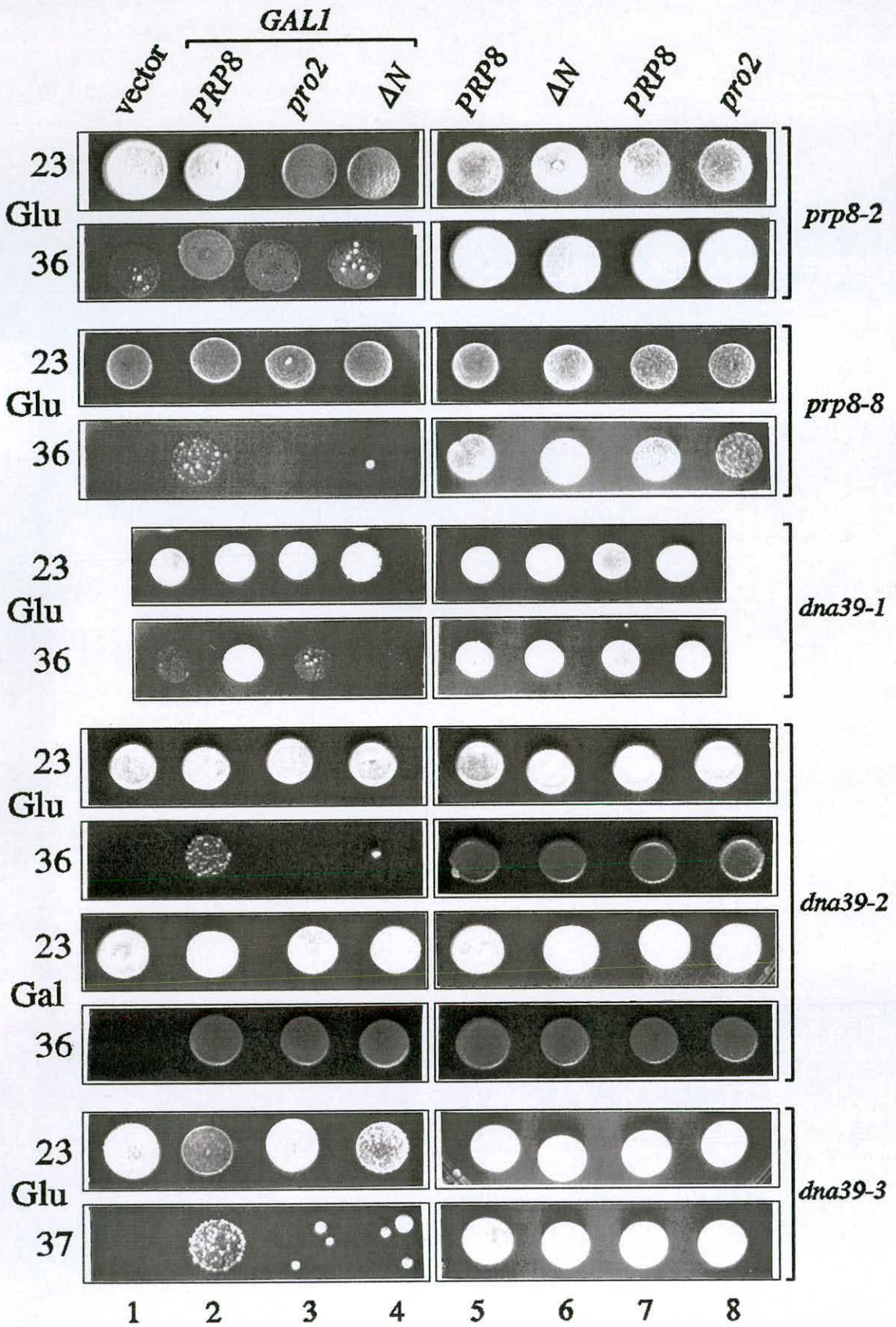

GAL1
vector
PRP8
pro2
ΔN
PRP8
ΔN
PRP8
pro2
23
Glu
36
prp8-2
23
Glu
36
prp8-8
23
Glu
36
dna39-1
23
Glu
36
23
Gal
36
dna39-2
23
Glu
37
dna39-3
1
2
3
4
5
6
7
8

Figure 7.4: Complementation by *prp8ΔN* is Linked to Pre-mRNA Stability.

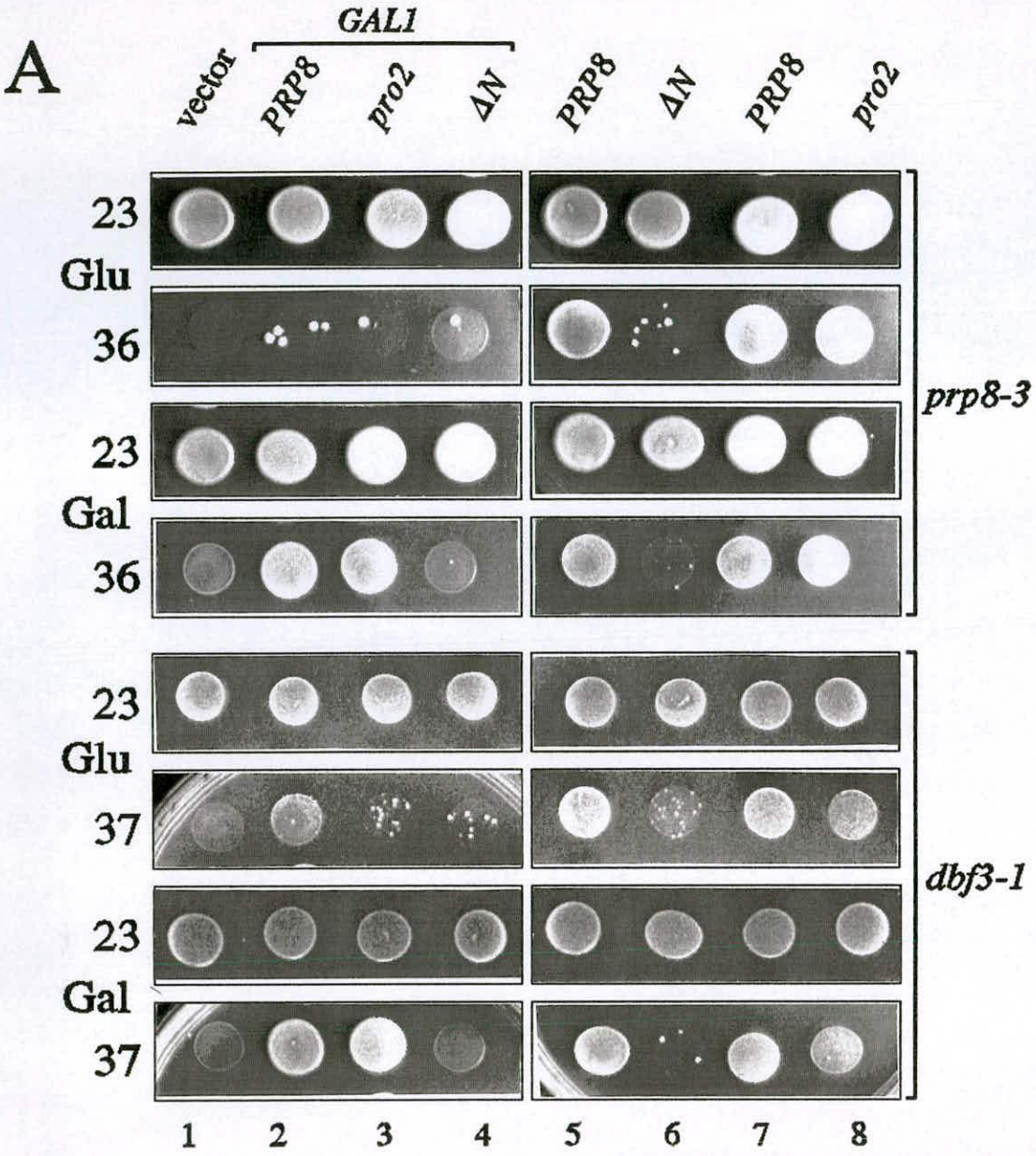
(A): A subset of prp8 strains (*dbf3-1, prp8-3, prp8-5*) is not complemented by *prp8ΔN*.

JDY8.31 (*prp8-3*) and L149-7B (*dbf3-1*) cells bearing URA3 plasmids were transferred to plasmid-selective plates for growth at permissive (23°C) or restrictive (36°C for *prp8-3* and *prp8-7*; 37°C for *dbf3-1*) temperatures under inducing (YMGRcas plates, Gal) or repressing (YMGlucas plates, Glu) conditions. Plates were incubated for 5 days. Data for JDY8.57 (*prp8-5*) transformants are not presented but temperature sensitivity of this strain is complemented on glucose likewise by pAJY13 (*prp8ΔP2*) but not by pJDY7 (*prp8ΔN*). The URA3 plasmids were:

(1) vector:	pBM125	(5) *PRP8*:	pJDY6
(2) *GAL1:PRP8*:	pBM-PRP8	(6) *prp8ΔN*:	pJDY7
(3) *GAL1:prp8ΔP2*:	pAJY1	(7) *PRP8*:	pAJY12
(4) *GAL1:prp8ΔN*:	pAJY9	(8) *prp8ΔP2*:	pAJY13

(B): Complementation of *dbf3-1* by *prp8ΔN* correlates with pre-mRNA stability.

AJY8.31, AJY8.32 and AJY8.33 (all *dbf3-1*, *ura3*) are haploid strains derived from outcrossing L149-7B into wild-type strain D150 (see Figure 7.2). Cells bearing *URA3* plasmids were transferred to plasmid-selective plates for growth at permissive (23°C) or restrictive (36°C) temperatures under inducing (YMGRcas plates, Gal) or repressing (YMGlucas plates, Glu) conditions. Plates were incubated for 5 days. Relevant features of the *URA3* plasmids are: pRS316 (*CEN*; vector), pY8000 (2μ, *PRP8*), others see above.


A
GAL1
vector
PRP8
pro2
ΔN
PRP8
ΔN
PRP8
pro2
23
Glu
36
23
Gal
36
prp8-3
23
Glu
37
23
Gal
37
dbf3-1
1
2
3
4
5
6
7
8

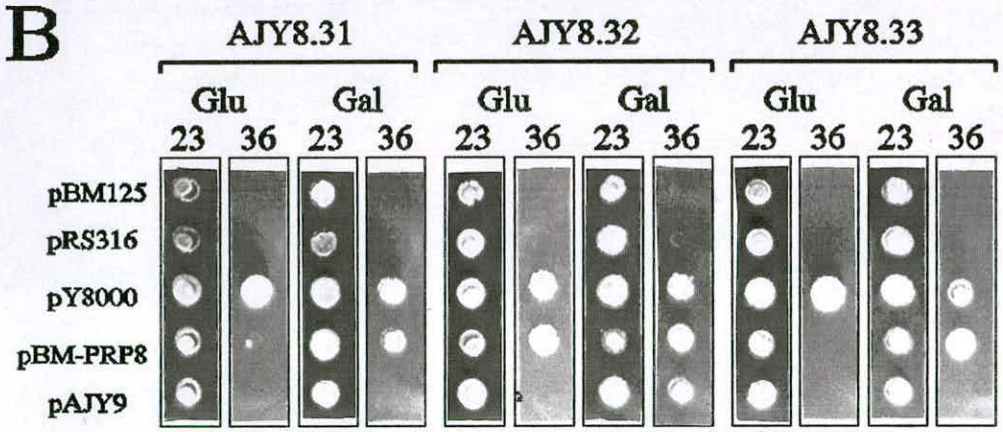

B
AJY8.31
AJY8.32
AJY8.33
Glu
Gal
Glu
Gal
Glu
Gal
23
36
23
36
23
36
23
36
23
36
23
36
pBM125
pRS316
pY8000
pBM-PRP8
pAJY9

Table 7.1

Correlation between Pre-mRNA Stability and Complementation by *prp8ΔN*.

Strain	Genotype	Complementation by *prp8ΔN*	Stability of accumulated pre-mRNA
SPJ8.31	*prp8-1*	✓	✓
JDY8.22	*prp8-2*	✓	✓
JDY8.31(I1)	*prp8-3*	X	X
JDY8.57	*prp8-5*	X	X
DJY76	*prp8-7*	✓	✓
JDY8.81	*prp8-8*	✓	✓
JL1U	*dna39-1*	✓	✓
JL2U	*dna39-2*	✓	✓
JL3U	*dna39-3*	✓	✓
L149-7B	*dbf3-1*	X	X
AJY8.31	*dbf3-1*	X	X
AJY8.32	*dbf3-1*	✓	✓
AJY8.33	*dbf3-1*	X	X

7.3: Complementation of *prp8-1* and *dbf3-1*.

The allelism *DBF3* and *PRP8* was unforseen because not only is the *dbf3-1* phenotype distinct from other splicing mutants, but *dbf3-1* and *prp8-1* complement each other. The complementation tests carried out in the characterisation of *dbf3-1* failed to reveal the identity of *DBF3* for this reason (Shea *et al.*, 1994). This intragenic complementation is independent of strain background, as several pairs of *prp8-1* and

dbf3-1 haploid strains form temperature-insensitive diploids, when crossed (Shea *et al.*, 1994). These *dbf3-1/prp8-1* diploids yield temperature-sensitive haploid progeny upon sporulation, excluding the possibility that recombination events restore a wild-type copy of the gene. Complementation is also specific to *prp8-1* and *dbf3-1*; no other pair of alleles which has been tested forms a diploid viable at 37°C (Shea *et al.*, 1994).

7.3.1: Introduction of *prp8-1* into Haploid L149-7B (*dbf3-1*).

Preliminary to studying this interesting behaviour, a plasmid-borne copy of *prp8-1* was introduced into haploid L149-7B (*dbf3-1*) in an attempt to reproduce complementation. It was found however that intragenic complementation could not be reproduced: L149-7B cells bearing pJDY23 (2μ, *prp8-1*) were still temperature sensitive at 37°C (Figure 7.5A). There are two key differences between this experiment and those described by Shea *et al.* (1994) and either could cause the discrepancy. One is that Shea *et al.* have demonstrated intragenic complementation only in diploids, whereas here the two mutations were combined in a haploid cells. The second is that *prp8-1* and *dbf3-1* were both chromosomal in the diploids created by Shea *et al.*, whereas here gene-dosage must differ markedly as *prp8-1* was on a high-copy plasmid (expected copy number, 20 to 50).

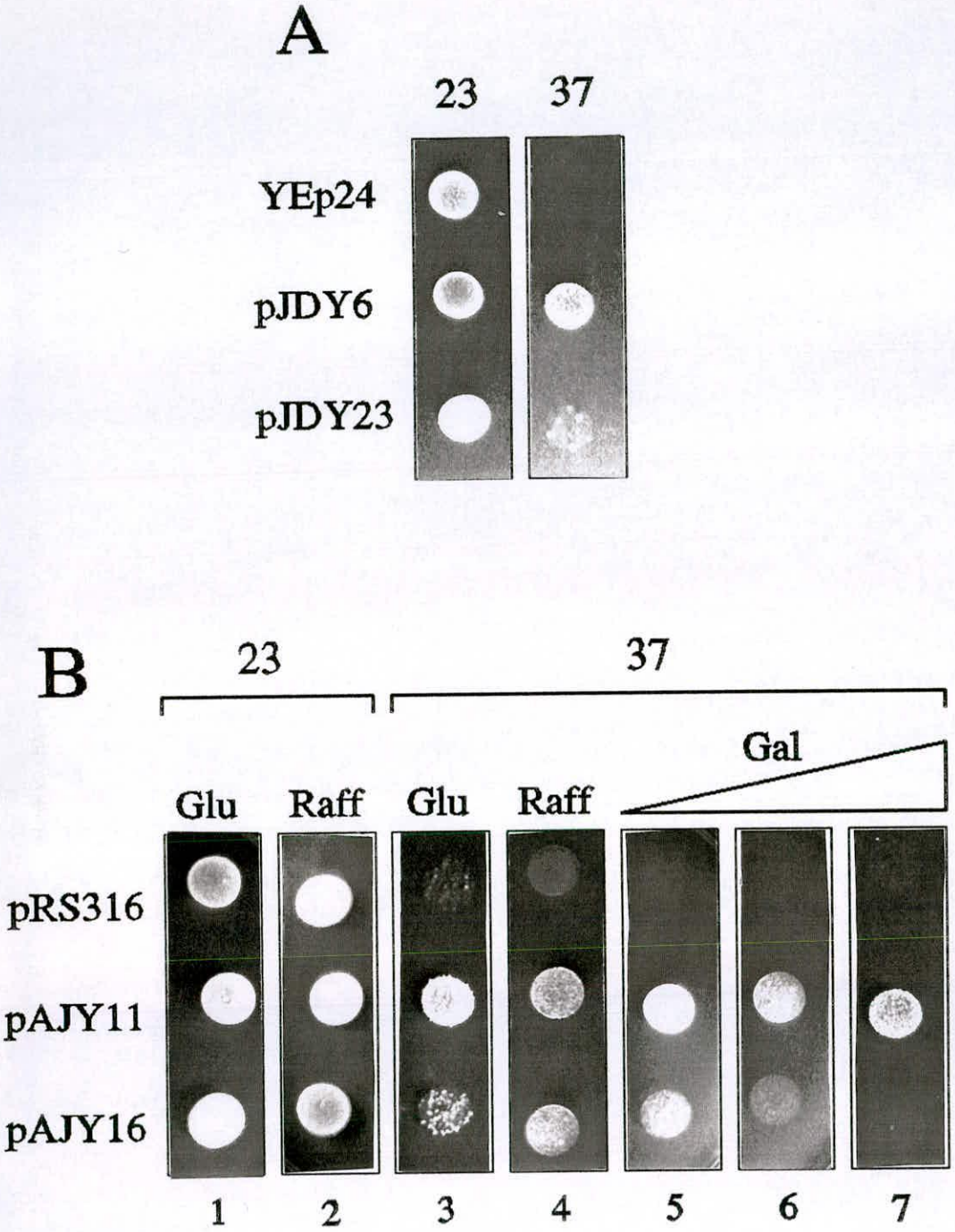
7.3.2: Modulation of *prp8-1* Expression by *GAL1:prp8-1* Fusion.

To determine if *dbf3-1* complementation could be reproduced by low-level expression of *prp8-1* expression, the *prp8-1* allele was transferred to the transcriptional control of the *GAL1* promoter of pAJY11 (*GAL1:PRP8*). Details of construction are given in the legend of Figure 7.5. Consistent with the experiment above, L149-7B (*dbf3-1*) cells bearing the *GAL1:prp8-1* construct pAJY16 were still temperature sensitive at 37°C on 2% galactose (YMGRcas plates, not shown). To modulate expression from the *GAL1* promoter, cells bearing pAJY16 were transferred to plates containing raffinose as a carbon source in combination with low concentrations of galactose for sub-maximal induction. Under these conditions yeast is expected to metabolise raffinose which is a superior carbon source, and not galactose which still causes *GAL1* induction but remains at a stable concentration. The alternative of mixing

Figure 7.5: Gene-dosage Dependence of *prp8-1/dbf3-1* Intragenic Complementation.

(A): *prp8-1* borne on a high-copy plasmid does not complement L149-7B (*dbf3-1*).
pJDY23 (*prp8-1*) was created by J.D.Brown (unpublished) and is identical to JDY6 (2μ, *PRP8*) except for the *prp8-1* point mutation (Hodges *et al.*, 1995). To show conclusively that the sequenced G to A mutation (Hodges *et al.*, 1995) in pJDY23 was responsible for the *prp8-1* phenotype, pJDY23 was isolated in the null background (JDY8.05; *prp8::LEU2*, pJDY13) by plasmid shuffle. The resulting strain was temperature-sensitive at 34°C (not presented). L149-7B (*dbf3-1*) cells bearing pJDY23 (2μ, *URA3, prp8-1*), pJDY6 (2μ, *URA3, PRP8*), or YEp24 (2μ, *URA3*) were resuspended and transferred to YMGlucas plates lacking uracil (supplemented with tryptophan). Plates were incubated for 7 days at 23°C or 37°C.

(B): Modulation of *prp8-1* expression to restore L149-7B complementation.
A plasmid containing *prp8-1* under the transcriptional control of the *GAL1* promoter (pAJY16) was created by replacing the *Cla*I fragment of pAJY11 (*GAL1:PRP8*) with the corresponding fragment, which encompasses the *prp8-1* point mutation, from pJDY23 (*prp8-1*). pAJY16 (*GAL1:prp8-1*) was isolated in the null background (JDY8.05; *prp8::LEU2*, pJDY13) by plasmid shuffle, and the resulting strain was galactose dependent, and temperature-sensitive on galactose at 34°C (not presented). L149-7B cells bearing pAJY11 (*CEN, URA3, GAL1:PRP8*), pAJY16 (*CEN, URA3, GAL1:prp8-1*), or pRS316 (*CEN, URA3*) were resuspended and transferred to minimal plates lacking uracil (supplemented with tryptophan). Plates contained as a source of carbon glucose (Glu; YMGcas plates; columns 1 and 3) or raffinose (Raff; YMGRcas except lacking galactose; columns 2, 4-7), and additionally galactose at 0.01% (column 5), 0.05% (column 6) or 0.1% (column 7). Plates were incubated for 7 days at 23°C or 37°C.


A
23
37
YEp24
pJDY6
pJDY23
B
23
37
Gal
Glu
Raff
Glu
Raff
pRS316
pAJY11
pAJY16
1
2
3
4
5
6
7

low concentrations of glucose with galactose is problematic because glucose is used preferentially as a carbon source, and represses *GAL1* transcription but changes concentration during the course of growth. As galactose was reduced below 0.1%, weak growth of pAJY16 transformants was detected, over and above background growth of control vector-transformed cells (Figure 7.5B). Maximal growth of L149-7B cells bearing pAJY16 was observed on 2% raffinose in the absence of any galactose (a non-induced, non-repressed state of intermediate expression of the *GAL1* promoter), but was still substantially less than that supported by expression of wild-type *PRP8* (from pAJY11). Possibly, reduction of *prp8-1* expression still further might complement *dbf3-1* more completely.

It therefore appears that *dbf3-1/prp8-1* intragenic complementation relies on matching levels of expression of the two alleles, and is "squelched" by overexpression of *prp8-1*, either due to *GAL1* overexpression or high gene-dosage. Complementation is not dependent on ploidy and can be reproduced in haploid as well as diploid strains.

7.4 Discussion.

The data presented in this chapter are consistent with those of previous chapters. The extent of glucose repression of pBM-PRP8 (*GAL1:PRP8*), defined by its ability to confer *PRP8* function, is variable between strains but in most cases glucose repression reduces expression to growth-limiting levels. In general, temperature-sensitive *prp8* strains are complemented by the truncation mutant *prp8ΔN*. This has been demonstrated with all available alleles except *prp8-3* and *prp8-5*, and the failure of complementation in these cases is most probably due to a strain background effect. Northern analysis with *prp8* alleles suggests that they all affect pre-mRNA splicing. The various splicing phenotypes of *dbf3-1/dna39* alleles (mRNA depletion with or without pre-mRNA accumulation) are typical of *prp8* alleles.

7.4.1 Precursor mRNA Instability.

The analysis of the range of *prp8* alleles revealed a subset which did not accumulate pre-mRNA, in spite of being associated with mRNA depletion at the

restrictive temperature. This effect is not caused by the particular *prp8* allele, but seems to affect any *prp8* allele when in a genetic background derived from DBY745. Steady-state pre-mRNA levels are increased by transcription and decreased by splicing and degradation. Where pre-mRNA is not accumulated in spite of a splicing defect, then logically either transcription must decrease or degradation must increase. It is remarkable that in SPJ8.31 and L149-7B strains grown at permissive temperatures the ratio between pre-mRNA and mRNA is approximately the same (compare lane 1 in Fig. 7.2A and lane 1 in Fig. 7.2B), even though one accumulates pre-mRNA at the restrictive temperature and the other does not. Possibly the factor which destabilises pre-mRNA in L149-7B does so only at the higher temperature. The experiment presented in Figure 6.7 indicates that the pre-mRNA accumulation caused by production of dominant negative $Prp2^{dn1}$ protein in the L149-7B strain was not sensitive to the destabilising effect. This is interpreted as indicating that production *in vivo* of toxic $Prp2^{dn1}$ protein causes stalling of fully assembled spliceosomes, as has been shown to occur *in vitro* (Plumpton *et al.*, 1994). Pre-mRNA may be able to accumulate in this situation because in the stalled spliceosome it is protected from the destabilising or degrading activity. This result does suggest that the absence of pre-mRNA accumulation in L149-7B due to *dbf3-1* is not due to changes in transcription.

Exactly what causes pre-mRNA instability in L149-7B (*dbf3-1*), JDY8.31(I1) (*prp8-1*), and JDY8.57 (*prp8-5*) is not known. Our original working model was that intron-containing pre-mRNA entered spliceosomes but was subject to aberrant splicing activity resulting in unstable products. For example, if step II utilised a non-authentic 3' splice site, an untranslatable and unstable mRNA might result. However, with the demonstration that the pre-mRNA stability phenotype was not associated with *prp8*, this model was reinterpreted. The *in vitro* characteristics of L149-7B (*dbf3-1*) extract (Chapter 6) also do not support this. Intron-containing pre-mRNA which is exported to the cytoplasm without being spliced is translated and becomes subject to degradation if in-frame stop codons occur in the intron. This is known as nonsense-mediated degradation (Peltz *et al.*, 1993, He *et al.*, 1993); its function is believed to be to prevent the production of truncated and potentially toxic polypeptides from mistakenly exported pre-mRNAs. One possibility is that in the subset of *prp8* strains pre-mRNA is more actively exported from the nucleus (or less well retained in the nucleus) than

usual, and degraded by this mechanism. Were this the case, a pre-mRNA which lacked stop codons would be expected to be accumulated as normal.

Screens for second-site mutations which are synthetic lethal in combination with *prp8ΔN* have been proposed as a means to identify factors which interact with the Prp8p N-terminal domain (Section 5.7.3). Pre-mRNA instability may be viewed as just such a mutation. To date, only combinations of temperature-sensitive and truncation mutants in pre-mRNA instability strains have been analysed, but it is expected from the results here that overexpression of *prp8ΔN* (in a strain containing no other *PRP8*) in the DBY745 background will not support growth. It is noted that other types of interaction as well as direct physical interaction can lead to a synthetic lethal phenotype. The product of the mutant allele which causes pre-mRNA instability might not be a spliceosomal factor at all but involved in RNA transport, or an hnRNP protein. Considering that the proposed N-terminal domain function is to catalyse steps of spliceosome assembly, the half-life of pre-mRNA may be crucial for the viability of truncation mutant strains. Any screen carried out in the future to identify further synthetic lethal interactions with N-terminal truncation mutants would be expected to yield the same type of second-site mutation characterised here; if necessary assessment of pre-mRNA instability may be used to exclude these from the analysis.

7.4.2 Intragenic *dbf3-1/prp8-1* Complementation.

Intragenic complementation between *dbf3-1* and *prp8-1* alleles seems to require equivalent expression of the two alleles. From the results presented here, it is predicted that *prp8-1* overexpression in *prp8-1/dbf3-1* diploids would confer a temperature-sensitive phenotype. Overexpression of *dbf3-1* in a *prp8-1* haploid, or in a *dbf3-1/prp8-1* diploid, might have similar effects. On the other hand, introducing *prp8-1* in single copy into a haploid *dbf3-1* strain (for example by chromosomal integration), is predicted to yield transformants viable at 37°C. These experiments could be readily carried out to corroborate (or otherwise) the results of modulating expression of a *GAL1:prp8-1* fusion presented here.

In the last chapter, evidence to suggest that *dbf3-1* has a mild-severity splicing defect was reviewed. The most obvious explanation of intragenic complementation

would be that summation of two partial splicing defects gives sufficient splicing activity to support growth, but this is unlikely because *prp8-1/prp8-1* and *dbf3-1/dbf3-1* diploids are still temperature-sensitive. The phenomenon of intragenic complementation often occurs in systems where the gene product functions as a multimer. For example the 25 kD product of the *RIB5* (riboflavin synthase) gene functions as a trimer, and truncation and point mutations can complement one another (Santos *et al.*, 1995). For similar examles see Carlson et al. (1981) and Friedman et al. (1994). As remarked previously, the stoichiometry of Prp8p in the U5 snRNP particle, and the U5 snRNP in the spliceosome is unknown. The possibility that each U5 snRNP contained one molecule of each mutants or that the spliceosome contains two U5 snRNPs cannot be discounted. Protein products of *dbf3-1* and *prp8-1* cannot currently be distinguished biochemically, as was possible with the truncation mutants (Chapter 4), but now that the *dbf3-1* mutation is sequenced (Chapter 6) and *prp8-1* reconstructed on plasmids differential epitope tagging of the two proteins is a reasonable prospect.The other common interpretation of intragenic complementation, is that the gene product exists as a monomer but has multiple independent functions. Each allele retains the ability to carry out one or some of these functions to confer viability to the heterozygote. Intragenic complementation amongst alleles of calmodulin are one such example (Ohya and Botstein, 1994), another is the *POL2* gene which encodes polymerase epsilon. Pol2p has an N-terminal polymerase domain and a C-terminal sonsor domain which detects blocked DNA replication acts as a checkpoint preventing progression to mitosis. Combinations of alleles mutant in the polymerase function with those mutant in the checkpoint function are viable (Navas *et al.*, 1995). Considering that in these examples the intragenic complementation phenotype has been highly informative of the action of the gene product, further study of *prp8/prp8ΔN*and *prp8-1/dbf3-1* heterozygotes is strongly warrated.

The requirement for equal expression of the two alleles may be due to competition. The effect of *prp8-1* overexpression in the *dbf3-1* strain may reduce expression from the chromosomal allele. *PRP8* contains numerous codons for rare tRNAs, and high levels of *PRP8* mRNA may effectively sequester these tRNAs and reduce protein production from the chromosomal *dbf3-1* allele. Alternatively, competition may be at a later stage such as entry into the U5 snRNP. If, as is suspected,

Prp8p is readily degraded unless assembled into ribonucleoprotein complexes, and if the products of the *prp8-1* and *dbf3-1* alleles compete for assembly into complexes, then the excess of protein from *prp8-1* might prevent assembly of the *dbf3-1* product, leaving it susceptible to degradation.

Overproduction of Prp8p from *prp8-1* either due to *GAL1* overexpression or high gene-copy, was not detrimental to growth of a wild-type strain, nor did it overcome the temperature-sensitive *prp8-1* defect. This contrasts with *prp2-1*, which is not temperature-sensitive when overexpressed (M. Lee, unpublished), and is consistent with the idea that Prp8p is not readily accumulated. It is difficult to predict what effect overexpression of other alleles such as *dbf3-1* would have on temperature sensitivity. If, as proposed, *dbf3-1* is a mild severity defect, overexpression of this allele might rescue temperature sensitivity.

Chapter 8.

Isolation of Suppressors of *dbf3-1* from a cDNA Expression Library.

8.1: Introduction.

The results presented in Chapter 6 identify *dbf3-1* as a *prp8* allele which affects splicing. This supports the hypothesis that the link between pre-mRNA splicing and the cell cycle stems from a particular intron-containing gene (or group of genes) whose expression is required for progression through S-phase (Shea *et al.*, 1994). Several known intron-containing yeast genes are involved in the cell cycle. *CDC28*, the primary regulator of cell cycle progression, lacks an intron but two other cyclin-dependent kinase (CDK) genes, *PHO85* and *KIN28*, are intron-containing. As part of a cyclin/kinase pair, Pho85p is periodically active in phase with the cell cycle and is most active in G1 (Measday *et al.*, 1994). Pho85p and two of its G1 cyclin partners, Pcl1p and Pcl2p, are each essential for progression through START under conditions where regulation of START is already compromised by deletion of the other G1 cyclins (Espinoza *et al.*, 1994). Kin28p is an essential CDK and is the functional homologue of metazoan MO15 (Feaver *et al.*, 1994). MO15 has roles in transcription and nucleotide excision repair as a component of transcription factor TFIIH, and also in cell cycle progression as the catalytic component of CDK-activating kinase (CAK) which phosphorylates other CDKs (including $p34^{cdc2/CDC28}$) at a conserved threonine residue which must be phosphorylated for CDK activity (Clarke, 1995). Overproduction of the cyclin partner of Kin28p, Ccl1p, can rescue a G1 cyclin triple-mutant (Cln1,2,3Δ).

Various transitions in the cell cycle involve specific proteolytic degradation. After cells pass START, the ubiquitin-conjugating enzyme Cdc34p targets the CDK inhibitor Sic1p to degradation in the proteasome, and Clb5p- and Clb6p-CDK complexes become active. Two of the four yeast ubiquitin structural genes have introns (*UBI1* and *UBI2*), as do at least three of the various known genes for ubiquitin-conjugating enzymes (*UBC4*, *UBC5* and *UBC8*). Furthermore, some of these have

deviant splicing signals: AAG, the 3' splice site of *UBC5*; CACUAAC, the branchpoint sequence of *UBC8*; and GUAUGC, the 5' splice site of *UBI1* (non-consensus nucleotides in bold). Thus these genes are candidates to link splicing and the cell cycle.

This chapter describes an attempt to isolate a gene linking splicing and the cell cycle, by screening a cDNA expression library for suppressors of *dbf3-1*. Conceptually, suppression would occur because expression of the cDNA would bypass the requirement for splicing in supply of the product of the intron-containing cell cycle gene. This approach is reliant on the supposition that the severity of the *dbf3-1* splicing defect is mild, such that other processes dependent on splicing (eg ribosome assembly) will continue sufficient to sustain growth after specific suppression of the cell cycle defect. Evidence concerning the relative severity of *dbf3-1* was discussed in Chapter 6. cDNAs from the primary screen whose expression conferred the ability to grow at 37°C were subjected to a secondary screen of Northern and partial sequence analysis to identify those derived from intron-containing genes.

8.2 Selection of cDNA Suppressors.

The cDNA expression library was constructed by Lui and co-workers (Lui *et al.*, 1992), and is based on the vector pRS316 (*CEN, ARS, URA3*). The library contains cDNA inserts expressed from the *GAL1* promoter. To isolate suppressor cDNAs which enabled the temperature-sensitive strain L149-7B (*ura3, dbf3-1*) to grow at 37°C, a direct selection approach was adopted. Library DNA was introduced into L149-7B (*dbf3-1*) and transformants spread to galactose-containing plates (YMGRcas lacking uracil and supplemented with tryptophan) to induce cDNA expression. Plates were incubated for 6 hours at room temperature to allow cDNA expression from the *GAL1* promoter to commence and then at 37°C, or 23°C for controls. Colonies which develop at 23°C have acquired the Ura^+ phenotype, and colonies which develop at 37°C have acquired both Ura^+ and temperature-insensitive phenotype. As a positive control L149-7B cells were transformed with pBM-PRP8 (*GAL1:PRP8*). With pBM-PRP8, selection at 37°C reduced transformation efficiency by approximately 50%, compared to selection at 23°C. This may be because recovery from transformation at 37°C requires induction of the *GAL1* promoter, whereas recovery at 23°C does not. The effect of

Table 8.1 Legend:

Introduction of the cDNA Expression Library into L149-7B (*dbf3-1*).

Temperature-sensitive L149-7B cells were transformed with pBM-PRP8 (approx. 1μg) according to the standard protocol. After transformation, cells were spread to minimal agar plates lacking uracil (YMGlucas (glucose) or YMGRcas (galactose) supplemented with tryptophan) to select for Ura^+ transformants. Plates were incubated at 23°C or 37°C (10% total transformed cells per plate; duplicate plates at each condition). Eleven similar transformations were carried out with plasmid DNA of the cDNA library (approx. 1μg per transformation). From each transformation, cells were spread to YMGRcas plates (10% transformed cells per plate) and incubated at 23°C (1 plate) or 37°C (9 plates). Selection of Ura^+ transformants on galactose reduced transformation efficiency almost 3-fold relative to glucose. Transformation efficiency on galactose (at 23°C) ranged from $2x10^3$ to $6x10^3$ colonies/μg. Selection for complementation of *dbf3-1* in addition to uracil prototrophy reduced transformation efficiency by approximately 50% on either medium, relative to selection for Ura^+ alone. The number of L149-7B transformants tested for suppression was calculated from 6450 x 99/11 (number of plates) x 0.5 (effect of temperature).

temperature on transformation of wild-type cells was not determined. Assuming that L149-7B cells transformed with a suppressing cDNA had a similar chance of recovering as cells transformed with pBM-PRP8, then roughly $3x10^4$ library-transformed L149-7B cells were screened for suppression (see Table 8.1).

Table 8.1

Transformation (plasmid DNA).	Condition.		Number of Plates.	Total Ura^+ colonies (all plates).
	Plate	Temp. (°C)		
pBM-PRP8 (*GAL1:PRP8*)	Glucose	23°C	2	1096
		37°C	2	576
	Galactose	23°C	2	406
		37°C	2	198
cDNA library (*GAL1:cDNA*)	Galactose	23°C	11	6450
		37°C	99	$2.9x10^4$ [estimate]

After transformation of L149-7B with the *GAL1*-cDNA library, a single colony was observed after 5 days incubation at 37°C. After a further 10 days incubation, 59 smaller colonies were observed. The fast growing colony was designated X, and the slow growing colonies numbered 1 to 59. Fast and slow growth rates may represent strong suppression, and weak, partial suppression respectively. A summary of the analysis of putative suppressors is presented in Table 8.2.

8.3: Identification of cDNA Suppressors.

8.3.1: Initial Analysis.

The 60 colonies obtained in the selection were resuspended in sterile water in wells of a sterile microtitre plate, replicated to YMGlucas plates either lacking or supplemented with tryptophan, and incubated at 23°C. All colonies were tryptophan auxotrophs, as expected for transformants of L149-7B, and demonstrating that the colonies obtained in the selection were not contaminants. To confirm that suppression was not due to chromosomal mutations, putative suppressor-containing colonies were replicated to YMGlucas and YMGRcas plates, and incubated at 37°C. All grew weakly on galactose (inducing conditions), but not at all on glucose (repressing conditions), showing that suppression was in all cases galactose-dependent and likely to be due to the plasmid-borne cDNA.

8.3.2 PCR Amplification of cDNA Inserts.

The next step in the analysis was to identify the putative suppressor cDNAs. Recovery of the library-derived plasmid to *E. coli* would have enabled restriction mapping and direct sequencing of the cDNA, and retransformation of L149-7B to show conclusively that suppression was linked to the plasmid. However, this procedure is impractical for such a large number of isolates. Instead, cDNA inserts were amplified by the polymerase chain reaction (PCR). This allowed rapid analysis of the large number of putative suppressors, but with the disadvantage that amplified cDNAs were not necessarily those responsible for the suppression. In particular, suppressing cDNAs of a length preventing efficient PCR amplification (>3-4kb) were excluded, and were likely to give rise to spurious results due to contamination with shorter, efficiently

amplified (but non-suppressing) cDNAs. Thus the analysis was a survey of potential suppressors, to reveal those deserving further study. Amplification does not prove that a cDNA possesses suppressing activity, which for each case requires plasmid recovery and retransformation of yeast.

cDNA inserts were amplified by PCR using primers G0897 and 762N which anneal adjacent to the cDNA cloning sites. Genomic DNA prepared from cultures of each suppressor colony in YMGlucas (23°C, maintaining selection for the plasmid) was used as template. Of the 60 PCR reactions, 48 (including X, the putative strong suppressor) yielded a product which in all cases was shorter than 2.5kb (see Table 8.2),.

8.3.3: Northern Analysis.

Northern blots of RNA extracted from *prp* mutant strains and probed for transcripts deriving from intron-containing genes yield characteristic patterns of bands. To determine if the cDNAs isolated in the screen derived from intron-containing genes, PCR products were radioactively labelled and hybridised to Northern blots. These blots were prepared from total RNA extracted from SPJ8.31 (*prp8-1*) cells grown at permissive (23°) and non-permissive (36°C) temperatures. Twenty PCR products were tested in this fashion, of which several (including X, the putative strong suppressor) yielded a hybridisation pattern suggestive of splicing: transcript species present at 23°C became depleted at 36°C. In some cases, accumulation of higher molecular weight species was detected concomitant to depletion of the shorter species (see Table 8.2). The Northern blots are not presented. Not all PCR products were tested, but the hybridisation patterns which were obtained suggested that the objective of isolating cDNAs from intron-containing genes had been fulfilled.

8.3.4: Partial Sequencing.

To determine if any of the cDNAs identified in the screen derived from previously identified genes with known involvement in the cell cycle, PCR products were sequenced. The primer (oligonucleotide G0879) annealed to *GAL1* to yield N-terminal sequence of the cDNA. In the majority of cases between 50bp and 200bp of sequence was obtained, which was compared to the GenEMBL database using the program FASTA. The results of these searches are presented in Table 8.2.

Table 8.2: Putative Suppressors of *dbf3-1*.

Column 1: Colonies with a phenotype of slow-growth obtained in the screen were numbered 1 to 59. The single putative strong suppressor was purified to a single colony before preparation of genomic DNA for PCR, and was designated X.

Column 2: F: Failure (yield of no product or multiple products).
Where the PCR product was used in Northern blotting/N-terminal sequence analysis, product size is indicated (note that some PCRs giving doublet products were processed for Northern blotting and sequencing). The PCR reaction (50µl) contained 2-5µg yeast genomic DNA. Cycling conditions (30 cycles) were 93°C, 30sec; 48°C, 30sec; 5min, 72°C.

Column 3: ND: Not determined.
Size of transcripts hybridising to labelled PCR product was estimated by comparison to positions of 5S, 18S and 23S rRNA, visualised by ethidium staining. Note that in several cases the PCR product is much shorter than the transcript to which it hybridises suggesting that in some clones (for example, #6), are of incomplete cDNA.

Column 4: ✓✓: Probable; ✓: Possible; X: Improbable.
Likelihood that cDNA is derived from an intron-containing gene, from the Northern signal.

Column 5: Two methods were used to sequence PCR products, either direct sequencing (I; Section 2.4.2.3) or by single primer reamplification (II, Section 2.4.2.2).

Column 6: pORF: Predicted open redoing frame.
Partial N-terminal cDNA sequence was compared to the GenEMBL database using the FASTA program (GCG7). cDNAs described as identical gave FASTA scores of at least 95% nucleotide sequence identity, and generally between 98 and 100%. Deviation from scores of 100% was due to ambiguities in the cDNA sequence. Because of these ambiguities, the potential of each N-terminal sequence to encode protein where no match was found in the database (ie presence of open reading frame) was not investigated. The N-terminal sequence data are included in Appendix B.

Table 8.2: Putative Suppressors of *dbf3-1*.

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
1	0.6	Putative mRNA (0.5) and pre-mRNA (0.9).	✓✓	I & II	*RPL32* (ribosomal protein gene; contains intron; same as #55).	
2	1.1	Transcript (1.0) depleted at 36°C.	✓	I	*RPL10* (ribosomal protein gene; no intron).	
3	0.8	Putative mRNA (0.6) and pre-mRNA (0.9).	✓✓	I & II	*BBC1* (a pORF on Chr XIII (cosmid 9375) with 47.4% identity to human breast basic conserved protein 1).	Potentially full-length cDNA.
4	F					
5	F					
6	1.3	Transcript (3.7) depleted at 36°C.	✓	II	Identity to 3.7kb uncharacterised pORF on Chr X (Acc. no. X88851).	Partial cDNA.
7	1.2	Transcript (1.1) unaffected by splicing defect.	X	I	*FBA1* (fructose bisphosphate aldolase; same as #8, #33).	
8	1.2	Transcript (1.2) unaffected by splicing defect.	X	I	*FBA1* (fructose bisphosphate aldolase; same as #7, #33).	
9	2.5	No Northern Signal Detected.		I	No Sequence Obtained.	

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
10	0.7	Transcript (2.0) unaffected by splicing defect.	X	I	High degree of homology (75-91%) to sugar transporter genes *HXT5*, *HXT6*, *HXT7* (hexose transporters), *GAL2* (galactose permease), *LGT3* (glucose permease).	Reifenberger *et al.* (1995).
11	F					
12	1.8	Transcript (2.1) depleted at 36°C.	✓	I & II	Identity to uncharacterised 1.4kb pORF on Chr IX (cosmid Sc8277). Contains putative intron.	Potentially full-length cDNA.
13	2.0	Transcript (2.6) depleted at 36°C.	✓	I & II	*CDC46* (Replication Licensing Factor, see Section 1.14; no intron).	cDNA includes ATG start codon.
14	1.3 1.8	Two transcripts (0.8, 0.9) depleted at 36°C, third transcript (1.0) accumulates.	✓✓	I	*PLC1* (gene involved in mitochondrial protein import; plant homologue is expressed specifically in S-phase).	Ito *et al.*, 1992.
15	F					
16	0.8	Transcript (0.75) unaffected by splicing defect.	X	I	Identity upstream of uncharacterised 0.4kb pORF on Chr VIII (cosmid 9986).	cDNA probably corresponds to 3' UTR.
17	1.6	Transcript (1.6) unaffected by	X	I	*TEF1* (transcription elongation factor;	Potentially full-

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
		splicing defect.			same as #43).	length cDNA.
18	0.6	Transcript (0.5) unaffected by splicing defect.	X	I	*RPS24* (ribosomal protein gene).	
19	1.0	No Northern signal.		I	Identity to uncharacterised 1.1kb pORF located on Chr III (position 303085 to 304191).	
20	1.4	No Northern signal.		I	66% identity over 130bp to human HMG-CoA synthase.	
21	1.4 1.5	Transcript (1.4) unaffected by splicing defect.	X	I	No Sequence obtained.	
22	1.2	ND		I	*GAP1* (glyceraldehyde phosphate dehydrogenase; same as #23, #48, #51).	
23	1.2	ND		I	*GAP1* (glyceraldehyde phosphate dehydrogenase; same as #22, #48, #51).	
24	1.0	Transcript (0.75) depleted at 36°C.	✓	I	Identity to uncharacterised 0.6kb pORF (designated YBR106w) on Chr II.	
25	1.8	ND		II	Identity to uncharacterised 0.3kb pORF on Chr VIII (cosmid 9332).	Adjacent to *NAM8* putative RNA binding protein.

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
26	2.1	Transcript (2.0) unaffected by splicing defect.	X	II	*SSB1* (hsp70 homologue; previously identified as a suppressor of *dbf3-1*; Shea *et al.*, 1994).	cDNA includes ATG start codon.
27	0.7	ND		II	*URP2* (probable component of the small ribosomal subunit).	Similarity to *X.laevis* rp S22 and S20.
28	F					
29	1.2 1.3	Transcript (1.0) unaffected by splicing defect.	X	II	No sequence obtained.	
30	0.75	ND		II	*RPLA4* (ribosomal protein gene).	
31	1.1	ND		II	Identity uncharacterised 0.8kb pORF on Chr V.	
32	1.5	ND		II	No match detected.	
33	1.4	ND		II	*FBA1* (fructose bisphosphate aldolase; same as #7, #8).	
34	1.3 1.4	ND		II	Identity to uncharacterised 1.1kb pORF on Chr IX (cosmid 9877).	
35	F					
36	1.85	No Northern signal.		II	Identity to uncharacterised 1.6kb pORF	

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
					(designated YBR214w) on Chr II.	
37	1.85	ND		II	*MET8* (encodes enzyme involved in methionine metabolism).	Cherest *et al.* (1990).
38	F					
39	F					
40	F					
41	1.0	ND		II	*ARO4* (DAHP-synthase).	
42	1.1	ND		II	*NAB1* (hnRNP protein; contains intron).	
43	0.7	ND		II	*TEF1* (transcription elongation factor; same as #17).	Partial cDNA.
44	0.8	ND		II	No match identified.	
45	1.8	ND		II	No match identified.	
46	0.9	ND		II	No sequence obtained.	
47	1.1	ND		II	*MPT4* (characterised as a suppressor of *POP2*, a gene encoding a proline-rich putative transcription factor involved in glucose derepression of gene expression).	Sakai *et al.* (1992).

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
48	1.2	ND		II	*GAP1* (glyceraldehyde phosphate dehydrogenase; same as #22, #23, #51).	
49	0.7	ND			*RPS13* (contains intron).	Region of cosmid Sc8337.
50	1.4	ND		II	No sequence obtained.	
51	1.4	ND		II	*GAP1* (glyceraldehyde phosphate dehydrogenase; same as #22, #23, #48).	
52	1.0	ND		II	*GCN4* (transcription factor).	
53	F					
54	1.6	ND		II	No sequence obtained.	
55	0.7	ND		II	*RPL32* (ribosomal protein gene; contains intron; same as #1).	
56	F					
57	0.8	ND		II	Identity to region downstream of *SSC1*. The region of identity is bounded by an unusual $(CA)_{20}$ block of sequence. Not pORF, may be 5'UTR of adjacent gene.	*SSC1* accession number is M27229.
58	0.7	ND				

#	PCR product (size in kb)	Signal on Northern Blot (size in kb)	Intron?	Method for sequencing	Result of FASTA search	Notes
59	F					
X	1.6	Transcript (1.7) depleted at 36°C.	✓	II	*TUB1* (α-tubulin gene; contains intron).	cDNA includes ATG start codon.

8.4 Weak Suppressors.

The N-terminal sequence data showed that one of the slow-growing colonies obtained in the screen contained a cDNA copy of the *SSB1* gene. *SSB1* encodes an hsp70 homologue, and does not contain an intron. Plasmid DNA was recovered from this transformant and introduced into *E.coli*; restriction mapping confirmed identity as *SSB1*. The plasmid was reintroduced into L149-7B (*dbf3-1*) and galactose-dependent suppression was reproduced. This gene is of interest because it was isolated previously in attempts to clone *DBF3* by complementation of thermosensitivity, and has been characterised as a high-copy suppressor of *dbf3-1* (Shea *et al.*, 1994). Its isolation in this study demonstrates the efficacy of the adopted approach, and is consistent with the estimation that a significant proportion of the library was screened.

Other slow-growing colonies contained cDNAs derived from *FBA1* (3 isolates), *GAP1* (4 isolates) and *TEF1* (2 isolates). These genes encode fructose 1,6-bisphosphate aldolase (FBA), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and elongation factor EF-1α, respectively, and none contains an intron. None of these plasmids was recovered to *E.coli* to reproduce the suppression, but their identification in independent transformants suggests that they do bear weak suppressing activity, and do not derive from contamination of PCR. It is possible that weak suppression of *dbf3-1* can be achieved by overexpression of any one of a variety genes, none of which has any obvious direct connection to pre-mRNA splicing. If indeed *dbf3-1* is associated with a mild splicing defect, the mechanism of suppression in these cases may be via global modulation of splicing level.

Several of the suppressors which were identified as deriving from intron-containing genes by Northern analysis encoded ribosomal proteins. Strictly, they do not have direct roles in the cell cycle; and they do not represent the type of intron-containing gene which the screen was intended to isolate. Again, as plasmids were not isolated to reproduce suppression, it cannot be discounted that rp cDNAs have been identified spuriously by contamination of PCR reactions: rp mRNAs are numerous and moderately abundant (Woolford and Warner, 1991). However, rp

cDNAs were identified in 5 of the 41 colonies from which sequence data were obtained, and one cDNA (*RPL32*) was isolated twice. The mechanism by which RP cDNAs might suppress a splicing mutant is not clear. Notably, the presence of an intron within an RP gene does not seem to have influenced its identification, as both intron-containing genes (such as *RPL32*) and intron-less genes (such as *RPL10*) were represented. cDNAs from the *UBI1* and *UBI2* genes, which contain introns and which encode rpS37 as well as ubiquitin (as a chimeric gene product; Woolford and Warner, 1991) were not obtained.

8.5: A Strong Suppressor.

The selection of cDNAs which confer on L149-7B (*dbf3-1*) the ability to grow at the non-permissive temperature generated one fast-growing putative suppressor colony. PCR amplified a unique product of approximately 1.6kb in length. Used as a probe this product identified a 1.7kb species in total RNA which was absent in RNA prepared from *prp8-1* cells after transfer to 36°C (not shown), consistent with the cDNA being derived from an intron-containing gene. N-terminal sequence identified the amplified cDNA as a copy of *TUB1*, a well-characterised yeast gene containing a single, typical yeast intron (Schatz *et al.*, 1986a and b). *TUB1* is one of two intron-containing yeast genes which encode alpha-tubulin, the other is *TUB3*. DNA was recovered and the plasmid introduced into *E.coli*. A single plasmid was recovered which had a restriction map characteristic of the published *TUB1* sequence; this plasmid was termed p*GAL1:TUB1*^{cDNA}. The *TUB1* cDNA begins at position -20 before the start codon, and the restriction map suggests that it extends to beyond the *TUB1* stop codon. It is therefore expected to encode full-length α-tubulin. L149-7B (*dbf3-1*) cells bearing p*GAL1:TUB1*^{cDNA} were temperature-sensitive on glucose, but grew at 37°C on galactose (Figure 8.1A). Thus, expression of a copy of the *TUB1* gene lacking an intron suppresses *dbf3-1*. As *TUB1* is involved in the cell cycle, encoding the structural component of the spindle enabling chromosome segregation in M-phase, it is potentially the link between splicing and the cell cycle.

Figure 8.1: Suppression of *dbf3-1* by Expression of *TUB1* and *SSB1* cDNAs.

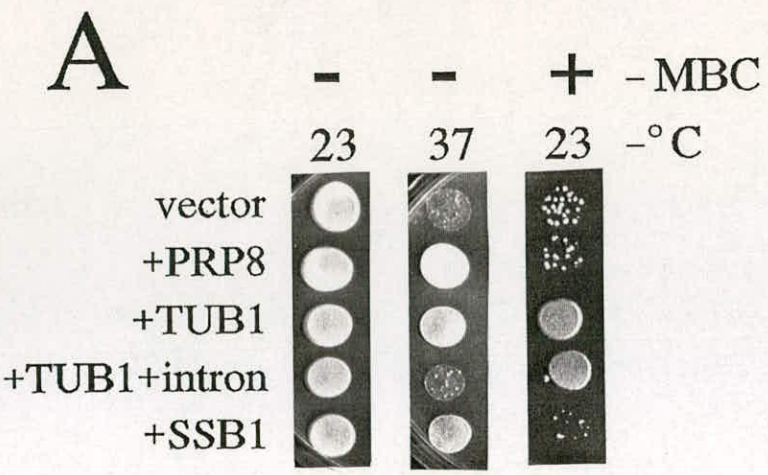
(A): Plate assay.

L149-7B (*dbf3-1*) cells bearing various *URA3* plasmids were resuspended in sterile distilled water to approximately 10^6 cells/ml, and transferred as 15µl spots to YMGRcas agar plates (inducing conditions) supplemented with tryptophan. Plates lacked uracil, to select for plasmid maintenance, and contained 1% DMSO, or 1% DMSO plus 25µg/ml methyl benzimidazol-2-yl carbamate (MBC). Plates were incubated at the permissive temperature (23°C) or the restrictive temperature (37°C) for *dbf3-1*. The *URA3* plasmids were:

Vector (row 1):	pRS316 (*CEN, ARS, URA3*)
+PRP8 (row 2):	pBM-PRP8 (*CEN, ARS, URA3, GAL1:PRP8*)
+TUB1 (row 3):	pGAL1:TUB1^{cDNA} (*CEN, ARS, URA3, GAL1:TUB1 cDNA*)
+TUB1+intron (row 4):	pGAL1:TUB1^{INT} (*CEN, ARS, URA3, GAL1:TUB1*)
+SSB1 (row 5):	pGAL1:SSB1 (*CEN, ARS, URA3, GAL1:SSB1 cDNA*)

(B): Expression of *TUB1* cDNA does not affect the *dbf3-1* splicing defect.

L149-7B (*dbf3-1*) cells bearing either pRS316 (lanes 1 and 2), pBM-PRP8 (lanes 3 and 4), or pGAL1:TUB1^{cDNA} (lanes 5 and 6; see above for characteristics of the plasmids) were cultured in YMGRcas liquid media supplemented with tryptophan to maintain selection for the plasmid, and induce *PRP8* or the *TUB1* cDNA. Cultures at 23°C were diluted two-fold into medium prewarmed to 23°C (lanes 1, 3 and 5) or 51° (lanes 2, 4 and 6), and incubation continued for 12 hours at 23°C and 37°C respectively. Total RNA prepared at this time was analysed by Northern blotting for *CYH2* transcripts as described previously (Figure 4.2C), and for *TUB1* transcripts by hybridising to a ^{32}P-labelled DNA fragment (1.6kb; *Bam*HI-*Sac*I, excised from pGAL1:TUB1^{cDNA}) encoding the entire *TUB1* gene.


A
- - + -MBC
23 37 23 -°C
vector
+PRP8
+TUB1
+TUB1+intron
+SSB1

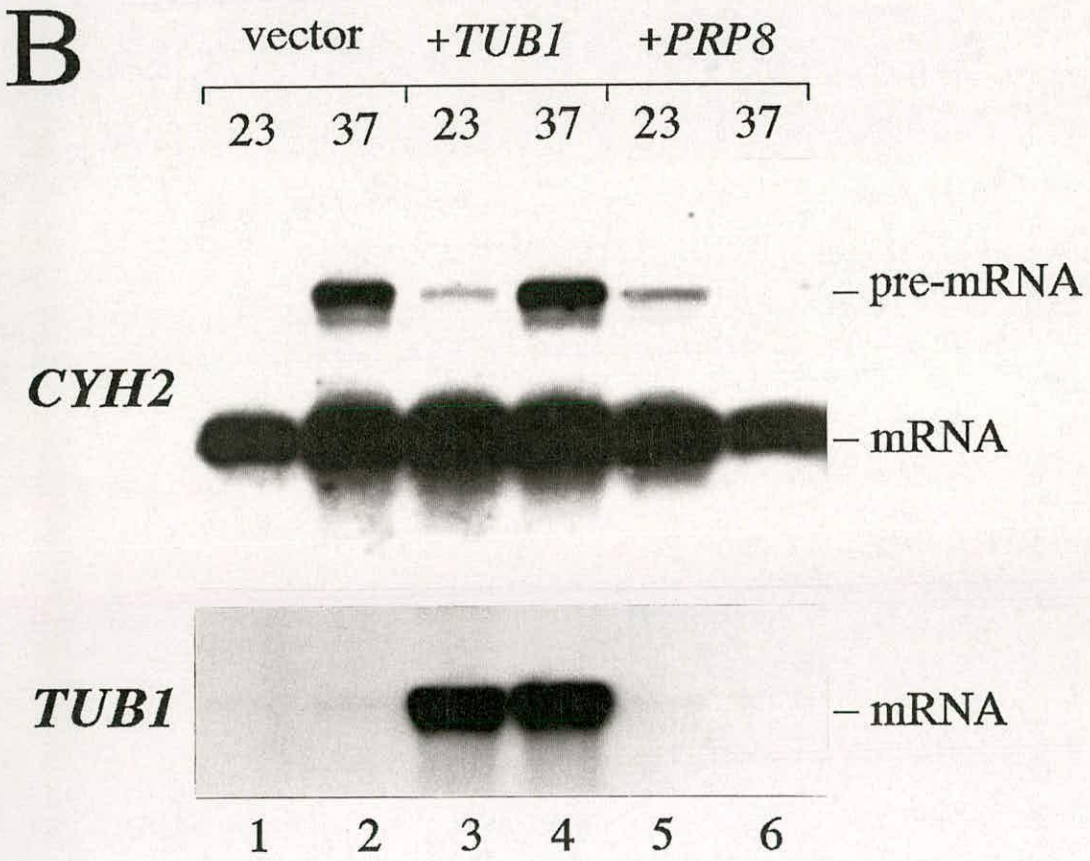
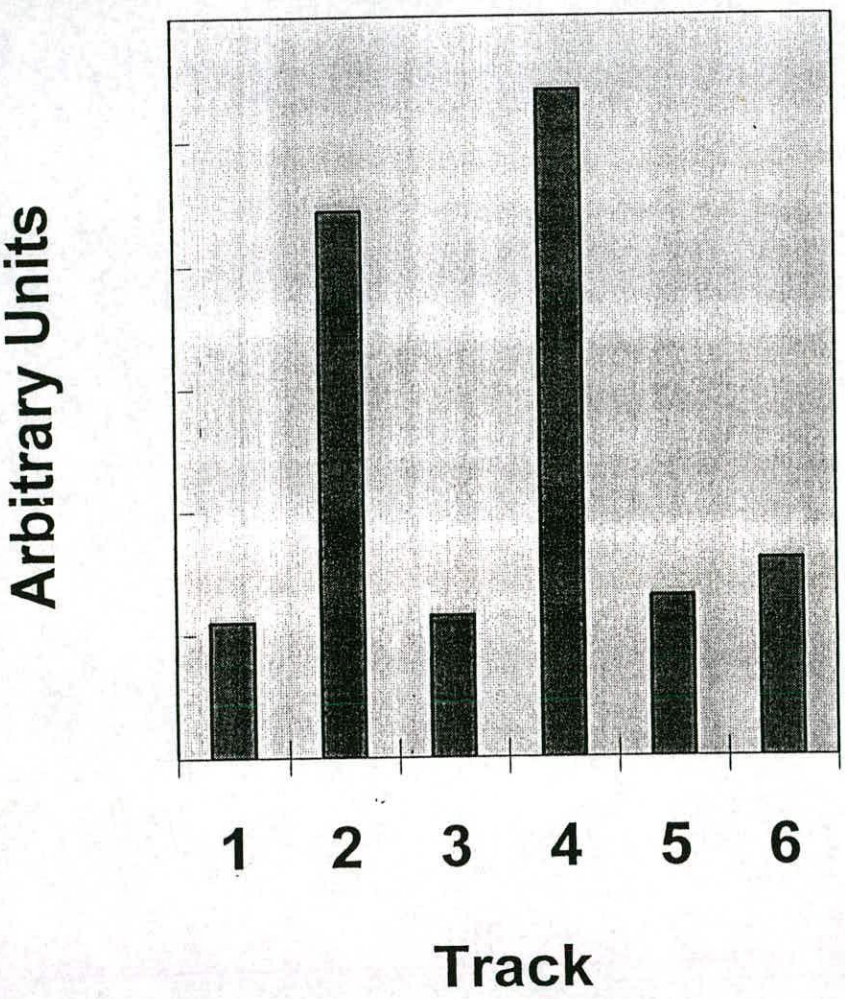

B
vector +TUB1 +PRP8
23 37 23 37 23 37
– pre-mRNA
CYH2
– mRNA
TUB1
– mRNA
1 2 3 4 5 6

Figure 8.2: Quantitation of Northern Blot (Figure 8.1B).

CYH2 mRNA and pre-mRNA detected in total RNA from L149-7B transformants (Northern blot presented in Figure 8.1B) were quantified using the Phosphorimager. The intensity of *CYH2* pre-mRNA in each track is expressed as a proportion of *CYH2* mRNA.

CYH2 pre-mRNA Level



However, its action in suppressing *dbf3-1* was unforeseen because treatment of yeast with drugs which prevent microtubule polymerisation arrests the cell-cycle at the beginning of M-phase, whereas *dbf3-1* cells reportedly arrest just after S-phase.

8.5.1: p*GAL1:TUB1*^{cDNA} Encodes Functional Tubulin Subunits.

Extra copies of functional α-tubulin genes alleviate sensitivity to microtubule-directed drugs, such as methyl benzimidazol-2-yl carbamate (MBC; Shatz, 1986b). To show that the suppressor conferred α-tubulin function, *dbf3-1* cells bearing p*GAL1:TUB1*^{cDNA} or the vector pRS316 were transferred to media containing various concentrations of MBC. *TUB1* expression from p*GAL1:TUB1*^{cDNA} enabled these cells to grow at 23°C at all MBC concentrations tested (up to 25 μg/ml; Figure 8.1A). When expression from p*GAL1:TUB1*^{cDNA} is prevented in the same transformants by glucose repression, concentrations of 5 μg/ml MBC and greater inhibited growth (not presented). L149-7B cells bearing the vector pRS316 are unable to grow on media containing high concentrations of MBC (Figure 8.1A).

8.5.2: Allele Specificity of Suppression.

To demonstrate that suppression was not influenced by strain background, strains AJY8.31 and AJY8.32 (both *dbf3-1*; see Section 7.2.1) were also tested for suppression. With both strains, cells bearing p*GAL1:TUB1*^{cDNA} grew on galactose but not glucose at the restrictive temperature. Thus suppression of *dbf3-1* does not seem to be influenced by the strain background. This also demonstrates that suppression by *TUB1* is independent of the pre-mRNA instability defect which was identified in L149-7B (and which is common to AJY8.31 but not AJY8.32). p*GAL1:TUB1*^{cDNA} was also introduced into the strains SPJ8.31 (*prp8-1*), JL1U (*dna39-1*), and JL3U (*dna39-3*). Suppression of these alleles could not be detected (not presented). In these strains the *GAL1* promoter is functional, at least sufficient to complement temperature sensitivity from the *GAL1:PRP8* plasmids and therefore the failure of suppression was probably not due to insufficient expression of *TUB1*. SPJ8.31 cells bearing p*GAL1:TUB1*^{cDNA} and grown under inducing conditions were resistant to anti-microtubule drugs, demonstrating that in this case functional α-tubulin was

being produced. Therefore *TUB1* cDNA is not an omnipotent suppressor of all *prp8* alleles, nor of the subset of *prp8* alleles with cell cycle defects, in fact suppression is specific to the *dbf3-1* defect.

8.5.3: Suppression is Specific to the Cell Cycle Defect.

At least two possible scenarios for the mechanism of suppression by overexpression of *TUB1* from p*GAL1:TUB1*^{cDNA} can be envisaged. One possibility is that *dbf3-1* cells are specifically deficient in α-tubulin, due to the inability to splice the *TUB1* and/or *TUB3* introns. Expression of *TUB1* cDNA bypasses the requirement for splicing in tubulin production, and the ability of cells to grow is restored. Alternatively, suppression may be indirect, for example by stimulating the general level of splicing in *dbf3-1* cells. To discount the latter possibility, Northern analysis was carried out. As L149-7B (*dbf3-1*) had been shown not to accumulate pre-mRNA, it had been intended to compare spliced mRNA level as a measure of splicing efficiency. However in this experiment (in which growth conditions differ from those used previously), some pre-mRNA accumulation was detected, and the pre-mRNA:mRNA ratio is presented (Figure 8.2). L149-7B (*dbf3-1*) cells bearing either the vector pRS316 (*CEN, URA3*), or pBM-PRP8 (*CEN, URA3, GAL1:PRP8*), or p*GAL1:TUB1*^{cDNA} (*CEN, URA3, GAL1:TUB1*) were incubated under inducing conditions (YMGRcas medium) at the restrictive temperature of 37°C for 12 hours, and total RNA prepared and analysed for transcripts of the *CYH2* and *TUB1* genes (Figure 8.1B). Cells bearing the vector pRS316 deviated from logarithmic growth, and ceased to grow beyond approximately nine hours. They had a splicing defect causing accumulation of *CYH2* pre-mRNA relative to mRNA (Figures 8.1B and 8.2). Cells bearing p*GAL1:TUB1*^{cDNA} contained high levels of *TUB1* mRNA due to expression from the plasmid (Figure 8.1B, lanes 3 and 4); in other lanes very faint signals were detected, corresponding to expression from the chromosomal *TUB1* (and possibly *TUB3*) locus. The cells suppressed by expression of *TUB1* cDNA continued to grow beyond the time that growth had ceased in the control, and the cultures reached stationary phase. These cells also had a splicing defect, quantitatively similar to that observed in controls. Expression of *PRP8* from pBM-

PRP8 rescued the growth defect and cells continued to grow logarithmically for many generations after transfer to 37°C. These cells contained approximately the same level of pre-mRNA (as a fraction of spliced mRNA; Figure 8.2) as the control grown at 23°C. Thus wild-type *PRP8* rescues both growth and splicing defects, whereas suppression mediated by the *TUB1* cDNA rescues the growth defect but not the splicing defect. In this experiment, growth conditions were the same as those used in the isolation of p*GAL1:TUB1*^{cDNA} as a suppressor (minimal media); Northern analysis following growth under these conditions has not previously been carried out. The fact that the cells accumulated pre-mRNA, whereas the same strain grown previously in complete medium (YPDA) was unable to do so, suggests that the factor which prevents pre-mRNA accumulation (Chapter 7) is variable according to growth conditions. Note that in this experiment the ratio of *CYH2* pre-mRNA to mRNA at 23°C is displaced towards the latter, compared to (for example) Figure 7.2. In this experiment, *dbf3-1* had very little impact on spliced *CYH2* mRNA level (Fig. 7.1B, compare lanes 1 and 2), even though previously measurable reductions have been observed. Again this is likely to stem from the differing growth conditions. It serves to reinforce the idea that *dbf3-1 in vivo* causes only a mild splicing defect. This experiment demonstrates that *TUB1* expression has no effect on overall splicing efficiency in L149-7B cells, consistent with the model that the mechanism of suppression is through bypass of splicing for a key cell cycle factor in which *dbf3-1* cells are deficient.

8.5.4: Insertion of the *TUB1* Intron into p*GAL1:TUB1*^{cDNA} Abolishes Suppression Activity.

To gain further support for the proposed mechanism of suppression, p*GAL1:TUB1*^{cDNA} was altered to restore the requirement for active splicing in production of α-tubulin. This was carried out by introducing the *TUB1* intron into the *TUB1* cDNA, to create a plasmid identical to p*GAL1:TUB1*^{cDNA}, except that the new plasmid expressed intron-containing *TUB1* pre-mRNA and not mRNA. An 0.4kb region of the chromosomal *TUB1* locus containing the intron was amplified by PCR with primers P0537 and P3333, using the plasmid pRB327 as template. The *Mlu*I-

*Xba*I region (280bp) of this PCR product was inserted into p*GAL1:TUB1*^{cDNA}, replacing the N-terminal end of the cDNA between the *Xba*I site (80 nucleotides downstream of the intron) and the polylinker *Mlu*I site adjacent to the cDNA cloning site. This construct, which was designated p*GAL1:TUB1*^{int}, contained 105bp of *TUB1* coding sequence (plus the intron) created by PCR. This region was not sequenced, and to exclude the possibility of PCR errors expression of functional tubulin was demonstrated by monitoring sensitivity to anti-microtubule drugs (Figure 8.1A). Both p*GAL1:TUB1*^{int} and p*GAL1:TUB1*^{cDNA} confer resistance to 25μg/ml MBC. Thus insertion of the intron did not interfere with the ability to express functional tubulin. However p*GAL1:TUB1*^{int} does not suppress *dbf3-1*, consistent with the hypothesis that in *dbf3-1* cells a primary deficiency is in α-tubulin, due to an inability to splice the *TUB1* intron. In effect, it is the *TUB1* intron which prevents division of *dbf3-1* cells.

8.6 Discussion.

8.6.1 Screen for Suppressors from a cDNA Expression Library.

The screen for cDNA suppressors of *dbf3-1* was successful, yielding a previously characterised high-copy suppressor of *dbf3-1* (*SSB1*), and one cDNA (*TUB1*) with the target characteristics of being a) derived from an intron-containing gene and b) essential for normal cell cycle progression. The estimate of transformation efficiency indicated that a reasonable proportion of cDNAs represented in the library were screened for *dbf3-1* suppression. Liu and co-workers recommended screening at least $5x10^4$ transformants for entire library coverage (Liu *et al.*, 1992), therefore this screen was not saturating, and some cDNAs particularly of rare transcripts will not have been tested. Notably, a cDNA copy of *PRP8* was not obtained, but the *PRP8* transcript is only weakly expressed (Jackson, 1988) and considering its length, is likely to be very poorly represented in the library. The cDNA library represents specifically the subset of RNA bearing a poly(A) tail, and does not include U snRNAs which cannot be identified in this screen.

Analysis of the collection of putative suppressors involved PCR and consequently it was expected that cDNAs greater than 4-5kb in length would be excluded. In the event, in all cases where a PCR product was generated, that product was never greater than 2.5kb in length. The most likely reason for failure of PCR for a fifth of all putative suppressors is that suppression was caused by longer cDNAs, which consequently have been eliminated from the analysis. In other cases, putative suppressor colonies may have contained long cDNAs with suppressing activity but short PCR products could have been produced, due to contamination. As the selection required simply growth at 37°C, it is possible that some colonies growing at 37°C contained two (or more) cDNA clones from the library. If both plasmids contributed to suppression, selective pressure would maintain the pair. In these cases, recovery to *E.coli* would be expected to generate mixed populations of transformants, and attempts to reproduce *dbf3-1* suppression would be expected to fail. It is notable that in several cases, PCR amplified doublets of products, and the presence of multiple cDNAs in the same transformant colony is a possible basis for this. From

these considerations, none of the cDNAs for which N-terminal sequence was obtained would wisely be studied in future as suppressors of *dbf3-1* without recovery of the plasmid to *E.coli*, analysis of *E.coli* transformants, and retransformation of L149-7B (*dbf3-1*) to reproduce suppression. In this study only those which have been recovered to *E.coli* can be regarded as being active as suppressors. Thus, the use of PCR was a limitation of this study, but was unavoidable due to time constraints. The collection of putative suppressors may contain interesting and informative cDNAs derived from splicing genes, cell cycle genes (intron-containing or otherwise) or others, and to enable future studies, cultures from all 59 putative weak suppressors have been stored.

8.6.2 Putative Weak Suppressors.

Multiple isolates of cDNAs encoding enzymes of glycolysis (GAPDH and FBA) and also cDNAs encoding elongation factor 1α were obtained in the screen (4 isolates of GAPDH, 3 of FBA, 2 of EF1α). The corresponding genes are strongly expressed in yeast, and thus expected to be well represented in the cDNA library. Multiple isolation does hint of activity in *dbf3-1* suppression, alternatively this may be the expected distribution of cDNAs (without *dbf3-1* suppressing activity) isolated at random from a cDNA library. It is interesting to note that precedent exists for relationships between splicing and glucose metabolism, and between splicing and translation. Specifically, mutations in *DED1* suppress *prp8-1*, and as already mentioned *DED1* is suspected to be involved in translational initiation (Jamieson *et al.*, 1991). Also, a mutation in the *SRN1* gene (also known as *HEX2/REG1*) whose product acts as a negative regulator of glucose repression can suppress a range of *prp* alleles including *prp8-1* (Pearson *et al.*, 1982, Tung *et al.*, 1992).

The isolation of the hsp70 homologue *SSB1* in both this screen and that of Shea *et al.* (1994), who screened a genomic library in an attempt to clone *DBF3*, confirms that this gene in high dosage suppresses *dbf3-1*, even though its relationship to pre-mRNA splicing is not obvious. Suppression by *SSB1* is specific to *dbf3-1* in that the *dna39* alleles are not suppressed, although the full range of *prp8* alleles has not yet been tested. Yeast contain numerous other hsp70 homologues (at least nine;

Craig *et al.* 1989) but in both this screen and that of Shea *et al.* (1994) only *SSB1* was identified, suggesting the activity is specific to this particular heat shock protein. A possible mechanism of suppression by *SSB1* is that the heat shock protein prevents the destabilisation of the association between Prp8p and the U5 snRNP particle, which accompanies heat-inactivation. The isolation of the *SSB1* cDNA in this study provides a useful internal control for the library screen.

8.6.3 Suppression by *dbf3-1* by Overexpression of *TUB1* cDNA.

None of the intron-containing genes with known or supposed S-phase functions proposed in Section 8.1 were identified in the screen. Instead, the familiar gene *TUB1* was identified. *TUB1* contains an intron and is essential for cell cycle progression, but is associated with M-phase rather than S-phase. However, the suppression of *dbf3-1* is more potent than with any of the other cDNAs identified, suggesting that the primary deficiency in *dbf3-1* cells is in α-tubulin. This is not the only defect: suppressed cells grow slower than cells complemented by *PRP8* (roughly 30-50%, judging by growth on plates). The partial nature of suppression is also clear from the observation that L149-7B (*dbf3-1*) cells bearing *PRP8* plasmids can grow at temperatures above 37°C, but suppressed cells cannot. Suppressed cells still have a splicing defect, and it is likely to be this that limits their growth. The splicing defect is consistent with a deficiency in a particular cell-cycle factor derived from a cell-cycle gene (namely TUB1) being responsible for growth arrest. In this model, suppression occurs because the cDNA bypasses the requirement for splicing in *TUB1* expression. A prediction of this was that an intron-containing TUB1 gene expressed under similar conditions as the cDNA would bear less potent suppressing activity. This was confirmed by constructing the plasmid p*GAL1:TUB1*^{int}.

Allele-specific suppression of *dbf3-1* by a cDNA copy of an intron-containing cell cycle gene casts was not expected in one of models proposed by Shea *et al.* (1994) that Prp8p and snRNP particles have dual roles, and function in cell cycle progression independently of their action in splicing. It is difficult to envisage how this hypothesis might be formally discounted, without constructing a yeast strain completely lacking in essential introns (an onerous if not impossible task). However,

there is no direct evidence to support this model, and that the results presented herein support the alternative, that the *dbf3-1* cell cycle progression phenotype is a consequence of a particular variety of splicing defect which causes aberrant production of a cell cycle protein.

8.6.4 Nature of the Defect in *dbf3-1* Cells.

The phenotypic effects of the *dna39* and *dbf3-1* alleles suggest a function specifically in S-phase. However, their phenotypes differ in an important respect: *dna39* cells arrest before S-phase with 1C DNA content, whereas in *dbf3-1* cells S-phase is delayed but does take place and arrest is in G2 phase with 2C DNA content (Shea *et al.,* 1994). The isolation of *TUB1* cDNA as a specific suppressor of *dbf3-1* further suggests that theses types of allele have distinct effects and that *dbf3-1* should be regarded as a G2 mutant and not an S-phase mutant as it was first described. G1 phase is main period of cell growth in the yeast cell cycle, and it is possible that S-phase is delayed in *dbf3-1* cells simply because the splicing defect reduces the rate of accumulation of cellular components. Other mutations in genes with no direct role in the cell cycle (for example, in components of the ribosome) have been observed to cause a similar S-phase delay.

Treatment of yeast with anti-microtubule drugs has the same end-result as *dbf3-1*, causing arrest in G2 phase. At a molecular level, however, the phenotypes are likely to differ. The drug treatments are thought to prevent polymerisation of α/β-tubulin dimers into microtubules, but probably do not affect the pool of α/β-tubulin dimers in the cell, or the ratio of α-tubulin to β-tubulin. This ratio is of crucial importance because excess β-tubulin is toxic (Burke *et al.*, 1989), causing growth arrest before nuclear division (M-phase). The molecular basis for this toxicity is not known. Thus it is expected that the expression levels of α- and β-tubulin are closely coupled. Apparently in wild-type cells α-tubulin is produced in slight excess to β-tubulin, and α-tubulin is degraded unless present in an α/β dimer. Overproduction of α-tubulin is not toxic, and the protein is not accumulated (Burke *et al.*, 1989). Therefore, α-tubulin deficiency may have the effect of releasing β-tubulin from the cellular pool of dimers, allowing it to have its toxic effect.

Another link between *dbf3-1* and S-phase is the execution point analysis described by Shea *et al.* (1994). It is remarkable that *dbf3-1* cells have a particular requirement for Prp8p activity during S-phase (as opposed to throughout the cell cycle), and again is most readily interpreted within the framework of a mild splicing defect. Our interpretation is that the "execution point" corresponds to the onset of a period of essential α-tubulin gene expression. Pre-mRNA generally has a much shorter half-life than mRNA, an active splicing apparatus would be required at the time of α-tubulin gene transcription to enable protein production. If transcription of tubulin genes was regulated during the cell cycle as is the case with other cell cycle genes, with transcription commencing around START and being active throughout S-phase, Prp8p activity would be required from the commencement of transcription and until sufficient α-tubulin had accumulated to sustain the subsequent M-phase. The promoter/enhancer elements which couple transcription to late G1/S phase are known; they are the SCB and MCB elements (Section 1.14.2). *TUB3* contains an SCB element (CACGAAA) in its upstream region, at positions -98 to -92 (relative to A of ATG start codon being +1). Exact matches to these elements in the *TUB1* upstream region have not been detected, but this region contains the sequence ACGCGAAA (at positions -155 to -148) which differs at only one position to the MCB consensus (non-consensus position in bold). This position of the MCB element has been demonstrated to tolerate change without abolition of transcriptional activation (McIntosh *et al.*, 1991). If indeed these elements couple *TUB1* and *TUB3* expression to S-phase, this may be why the product of *dbf3-1* can be heat-inactivated with consequences for cell cycle progression only during a narrow window coincident with S-phase. Note that positions -155 to -148 in the *TUB1* upstream region could alternatively be regarded as a SCB element (CGCGAAA), again differing at one position from the consensus.

One of the putative weak suppressors identified in the screen, a cDNA copy of *CDC46*, may also link *dbf3-1* and S-phase, as this gene encodes one of the *MCM* family of replication licensing factors (see Section 1.14.3) which are essential for DNA replication and whose function is to couple S-phase to M-phase so that DNA replication occurs only once during the cell cycle. Again it is stressed that the *CDC46*

cDNA was not recovered and reintroduced into L149-7B (*dbf3-1*) and should not be regarded as a true suppressor until this has been carried out.

8.6.5 Is *dbf3-1* Specific for the *TUB1* Intron?

From these results, it is not possible to conclude that the *TUB1* intron is spliced particularly poorly in *dbf3-1* cells. The cell cycle phenotype may result from a (mild-severity) splicing defect which does not differentiate between introns. The apparent difference between *TUB1* and other introns may be that of all products of intron-containing genes, α-tubulin is closest to being limiting for growth. However, if this were true it is surprising that overexpression from p*GAL1:TUB1*^{int} did not suppress *dbf3-1*. If any mild splicing defect resulted in such a clear block in cell cycle progression, then a defect whose severity could be varied (for example, different expression levels of a dominant negative protein) would portray a conditional cell cycle defect, and this has never been detected. Also, *prp* mutations would regularly be identified in screens for cell cycle mutants, and to date only mutations in *PRP3* and *PRP8* have been identified in this way (although perhaps if screens for cell cycle mutants routinely employed complementation analysis to the range of *prp* alleles, more splicing defects resulting in cell cycle phenotypes would be known). Lastly, if the cell cycle phenotype is a consequence of the unusual intron distribution in *S. cerevisiae*, why have splicing mutants been revealed from cell cycle screens in *S. pombe*?

These considerations appear to support the alternative notion, that *dbf3-1* splices the *TUB1* intron particularly poorly. The *TUB1* intron is typical for yeast: short (115 nucleotides), close to the 5' end of the pre-mRNA, and with consensus 5' and 3' splice sites and branchpoint. Thus it contains no obvious features to distinguish it for differential splicing. Northern blots of total RNA from *dbf3-1* cells probed for *TUB1* (not presented) clearly show depletion of spliced mRNA, but not a noticeably more severe splicing defect than for other intron-containing genes such as *RP28*. Another *prp8* mutant, allele *prp8-101* (Umen and Guthrie, 1995) has differential effects on the splicing of different introns, in that introns which require polypyrimidine tract recognition are inefficiently spliced. This set includes *TUB3*

which like *TUB1* encodes functional α-tubulin but which possesses an intron with an atypically long branchpoint to 3' splice site distance. In wild-type cells *TUB3* is expressed at much lower levels than *TUB1* and *TUB3* is dispensable whereas *TUB1* is essential, even though they are functionally interchangeable under conditions of ectopic expression. This explains the lack of effect on growth rate of *prp8-101*. However *TUB3* may be relevant if in *dbf3-1* cells *TUB1* splicing is compromised and *TUB3* splicing completely abolished, and hence the effect on α-tubulin levels exacerbated. It is feasible that *dbf3-1* affects polypyrimidine tract recognition in the same way as *prp8-101*, and would be readily investigated, using the reporter plasmids constructed in the study of *prp8-101* (Umen and Guthrie, 1995). Alternatively, *dbf3-1* may splice *TUB1* poorly for some other reason. No firm conclusions regarding the intron specificity of *dbf3-1* can currently be made.

8.6.6 (Lack of) Suppression of Other Alleles

The *TUB1* cDNA plasmid p*GAL1:TUB1*^{cDNA} did not suppress the *prp8-1* allele, which portrays no cell cycle progression phenotype; this was expected. An objective of this screen was to isolate an omnipotent suppressor active on all splicing mutations with cell cycle defects (or at least the other *prp8* mutations). However, the *TUB1* cDNA is not such an omnipotent suppressor, as it is unable to suppress *dna39* alleles, and this objective remains unfulfilled. The absence of *dna39* suppression by *TUB1* confirms that the *dna39* phenotypes are distinct from *dbf3-1*, a characteristic which was clear from the cell cycle analysis (Shea *et al.*, 1994; L.Johnston, personal communication). Incubation of *dna39* cells at the restrictive temperature is lethal and the cells arrest with 1C DNA content and have an earlier execution point. As far as splicing is concerned, two scenarios must be entertained. Firstly, it is possible that there exists another intron-containing cell cycle gene (not *TUB1*) which links the *dna39* alleles and the cell cycle. Suppression from a cDNA expression library would be a possible means to isolate that gene. Secondly, the cell cycle block in *dna39* cells may result from α-tubulin deficiency, but cells may be unable to grow even with *TUB1* cDNA due to a more severe splicing defect than *dbf3-1*. This hypothesis could be tested from the prediction that *TUB1* cDNA expression in *dna39* cells would

uncouple the cell cycle defect, and cause cells to arrest at random points of the cell cycle . As tubulin has no known essential functions in G1 phase however, the latter scenario is unlikely.

Also, it may be worthwhile to test *dbf5-2* and *dbf5-3* for suppression by *TUB1* cDNA. These are alleles of *PRP3* isolated at the same time as *dbf3-1*. They require to be outcrossed into transformation-proficient strains. Cells of *dbf5* have the same terminal morphology as *dbf3-1*, and which is not shared by *dna39* cells; *TUB1* cDNA may be an omnipotent suppressor of dumbbell-forming splicing mutants. Any demonstration of suppression of alleles other than *dbf3-1* would corroborate the data presented here.

Chapter 9.

Final Discussion.

At the outset of this work, Prp8p was familiar as a component of the yeast splicing apparatus with a central role in pre-mRNA splicing. However the Prp8 protein sequence was unremarkable for most of its length, revealing no significant homologies with other published proteins. Several conditional alleles had been characterised, but these seemed to have similar phenotypic effects (Brown, 1992). Thus the subdivision of this large protein into domains on the basis of function was not possible, and was an objective of this research.

The mutational analysis presented in this thesis has focused on two regions of Prp8p, to highlight putative or possible functions of each. The first region is the N-terminal domain, the most striking sequence feature of Prp8p, with four tandemly repeated copies of the LP_nG (n=5 to 8) peptide motif and acidic nature. The domain appears to have an accessory function, consistent with its absence from Prp8p homologues of higher metazoans, which are otherwise extremely well conserved. The second region corresponds to the location of a conditional point mutation, *dbf3-1*, which has been mapped to a central region of *PRP8* and sequenced. This mutation is of particular interest because it affects the cell cycle; under restrictive conditions *dbf3-1* cells arrest in the G2 phase with replicated chromosomes but prior to mitosis. The phenotype is highly unusual for a gene involved in pre-mRNA splicing. The initial analysis determined the effects of *dbf3-1* on splicing, which had not previously been characterised. This led to a screen for suppressors of *dbf3-1* from a cDNA expression library, and support for a hypothetical cause of the cell cycle phenotype.

The N-terminal domain of Prp8p is dispensable under certain conditions. For example, this region is not required if the truncated mutant lacking the domain is overexpressed. In this situation, although the cells are viable, defects in both growth rates and splicing activities are manifested. The yeast overexpressing the truncated mutant *prp8ΔN* yield reduced quantities of the snRNP complex U4/U6.U5 in an immunoprecipitation analysis. The requirement for overexpression implies roles for

the N-terminal domain in earlier assembly steps also. Taken together, the *in vivo* and *in vitro* phenotypes suggest that the role of this domain is to promote assembly of splicing complexes. The domain is also non-essential when a heat-inactivated version of Prp8p is present in the same cell. In this case *prp8ΔN* does not require to be overexpressed. The most plausible explanation of this is that the N-terminal domain can act *in trans*, perhaps promoting early steps of U5 snRNP assembly without being a component of that snRNP. It would be interesting to determine in future if this intragenic complementation requires full-length Prp8p, or if a fragment of a *prp8* allele will suffice.

The Prp8p N-terminal domain exhibits several features typical of proline-rich protein regions. An analysis similar to that carried out on transcription factors (Gerber *et al.*, 1994) has shown that a single copy of the LP_nG motif at the Prp8p N-terminus is sufficient to confer function (when n=10). No acidic residues are required, hence proline seems to be the most important feature. Gerber *et al.* found that homopolymeric stretches of proline function as activation domains, so possibly the functions of these domains are analogous. The action of transcription activation domains is still not fully clear, but in general they are thought to enhance the affinity of factors involved in transcription for the promoter region. For example, RNA polymerase might interact cooperatively with both the activation domain of a transcription factor and the basal transcription complex before initiating transcription. The chemical properties of proline are ideally suited to provide a non-specific but reasonably high-affinity interaction.

An obvious future direction of this research would be to determine the associative partner(s) of the N-terminal domain important in its functional role in the spliceosome. Typically, the 2-hybrid strategy to screen a library fuses the gene-of-interest to a DNA-binding domain; in the case of the Prp8p N-terminal domain such a fusion is itself expected to activate transcription from the sequence similarity to known activation domains. An alternative might be to fuse the N-terminal domain to an activation domain, and to combine this systematically with known splicing factor genes fused to DNA-binding domains. Another common approach is to isolate from a pool of randomly generated mutants those which are synthetic lethal with mutations

in the gene-of-interest. With *prp8ΔN* random screening may not be necessary because a feature of the genetic background of certain strains, presumably caused by a mutation, exhibits synthetic lethality with *prp8ΔN*. This background or mutation is amenable to standard genetical techniques for further analysis. There is some concern that the particular factor which is mutated is not a component of the spliceosome, considering it also apparently affects pre-mRNA stability, and therefore possibly does not interact physically with the N-terminal domain. This is still a potentially fruitful line of enquiry.

The *dbf3-1* allele was of great interest because the causal relationship between the mutant splicing factor and defective cell cycle progression lacked explanation. The workers who first characterised *dbf3-1* (Shea *et al.*, 1994) proposed two alternative models: that the splicing defect affected the expression of particular cell cycle genes, or that splicing factors had secondary and unrelated functions, in this case in the cell cycle. The second model, if true, would force the reinterpretation of much prior study. The first model implies that *dbf3-1* strains can splice some introns better than others, and thus offered the opportunity to study a potential role in intron recognition. The data presented in this thesis raise as many questions as they answer, and this topic is by no means closed, but broad support is provided for the first of the two models above. They demonstrate that splicing is inhibited at restrictive temperatures in a *dbf3-1* strain. *In vitro*, *dbf3-1* causes a near-complete inhibition of splicing, and the inactive Prp8p protein is poorly able to assemble with snRNA in snRNP complexes. *In vivo*, however, there is good evidence that *dbf3-1* does not completely abolish splicing but sustains a reduced level. Therefore the direct extrapolation of *in vitro* assays to *in vivo* activities is inappropriate. The *dbf3-1* allele was suppressed in an allele-specific fashion by a cDNA (ie intron-less) copy of a cell cycle gene. The particular gene is *TUB1*, one of two genes which encode α-tubulin in yeast. Tubulins are subunits of microtubules, and this explains neatly why *dbf3-1* cells arrest in G2. Suppression by the *TUB1* cDNA does not measurably affect the splicing of intron-containing transcripts, and is abolished by inserting the *TUB1* intron into the cDNA construct. These characteristics are consistent with the bypass-

of-splicing model: *dbf3-1* is deficient in α-tubulin, and the cDNA circumvents the requirement for splicing in α-tubulin expression.

Another observation of Shea *et al.* (1994) was of intragenic complementation between *dbf3-1* and *prp8-1* alleles. In the work presented here this finding is corroborated by expressing episomal *prp8-1* in the *dbf3-1* strain. The deficiency in *dbf3-1* cells seems to be in the level of α-tubulin, and the basis of complementation could be that *prp8-1* restores *TUB1* splicing. Even so the mechanism of intragenic complementation at the molecular level and its specificity to this pair of alleles is mysterious. Previously, *prp8-1* was thought to be unable to enter spliceosomes, and this may be another situation where the direct extrapolation of *in vitro* analyses to behaviour *in vivo* is not wholly appropriate. In the case of calmodulin mutants the study of intragenic complementation has been highly informative of the roles of the wild-type protein; the same will hopefully be true of Prp8p.

Conclusive evidence of differential splicing due to *dbf3-1* has not yet been presented. It is possible that any mutation which does not completely abolish splicing but which supports a residual degree (as does *dbf3-1*) also leads to tubulin deficiency and G2 arrest. However the rarity of alleles of splicing genes revealed in work on the cell cycle supports the differential splicing idea. Intriguingly, another *prp8* mutant, has also been linked to α-tubulin genes. This allele (*prp8-101*) can splice most introns tolerably well but not that of *TUB3*, the second yeast gene encoding α-tubulin (Umen and Guthrie, 1995). The splicing defect of *prp8-101* is specific to introns with long polypyrimidine tracts and has been demonstrated by direct assay. The application of the same assays to *dbf3-1* would readily determine if it affects splicing similarly. If the two mutations do have similar differential splicing effects, the positions of the two mutations may correspond to regions of the Prp8p polypeptide which assemble into a single domain in the native structure. Prp8p exerts direct contact on the polypyrimidine tract during the splicing (Teigelkamp *et al.*, 1995), so the domain of Prp8p disrupted by *prp8-101* (and possibly *dbf3-1*) may have an RNA-binding function. In the absence of robust biochemical methods to determine which region(s) of Prp8p contacts the pre-mRNA, the study of mutations with phenotypes like that of *prp8-101* is a vital complement to cross-linking analyses. The mutational

analysis has the advantage of revealing the importance of a particular interaction, which from cross-linking alone can be difficult to assess.

In conclusion, this work has started to delineate functional domains of Prp8p, although much work remains to be done. In the process, the mystery of why *dbf3-1* affects the cell cycle has been addressed. It is hoped that in the future the structure of Prp8p will become clear in ever-increasing detail, in parrallel with knowledge of the RNA structures which are thought to line the spliceosome active site. Of course the two are interlinked: to describe one is to describe the other. The role of genetics in this process will doubtless remain, for the spliceosome and its active site are highly complex and dynamic structures and will be difficult to solve by structural analyses alone. They will require the added power of the combination of biochemistry with genetics.

References.

Abovich, N., Legrain, P., and Rosbach, M. (1990) Mol. Cell. Biol. **10**:6417-6425.

Abovich, N., Liao, X.C., and Rosbach, M. (1994) Genes Dev. **8**:843-854.

Adzhubei, A.A., and Sternberg, M.J.E. (1993) J. Mol. Biol. **229**:472-493.

Aebi, M., Hornig, H., Padgett R.A., Reiser, J., and Weissmann, C. (1986) Cell **47**:555-565.

Aebi, M., Hornig, H., and Weissmann, C. (1987) Cell **50**:237-246.

Anderson, G.J., Bach, M., Lührmann, R., and Beggs, J.D. (1989) Nature **342**:819-821.

Anderson, G.J. (1989) PhD Thesis, Edinburgh University.

Antoniou, M. (1995) *In* Lamond, A.(ed.) Pre-mRNA processing. Springer-Verlag, Berlin. p. 187-195.

Arenas, J.E., and Abelson, J.N. (1993) Proc. Natl Acad. Sci. USA **90**:6771-6775.

Augustin, S., Muller, M.W., and Schweyen, R.J. (1990) Nature **343**:383-386.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K. (1987) Current Protocols in Molecular Biology. John Wiley and Sons, Inc.

Barabino, S.M.L., Blencowe, B.J., Ryder, U., Sproat, B.S., and Lamond, A.I. (1990) Cell **63**:293-302.

Beggs, J.D., van den Berg, J., Ooyen, A., and Weissmann, C. (1980) Nature **283**:835-840.

Beggs, J.D. (1995) *In* A. Lamond (ed.) Pre-mRNA processing. Springer-Verlag, Berlin. p. 79-95.

Behrens, S.-E., and Lührmann, R. (1991) Genes Dev. **5**:1439-1452.

Behrens, S.-E., Tyc, K., Kastner, B., Reichelt, J, and Lührmann, R. (1993) Mol. Cell. Biol. **13**:307-319.

Bennet, M., Pinol-Roma, S., Staknis, D., Dreyfuss, G., and Reed, R. (1992a) Mol. Cell. Biol. **12**:3165-3175.

Bennet, M., Michaud, S., Kingston, J., and Reed, R. (1992b) Genes Dev. **6**:1986-2000.

Bennet, M., and Reed, R. (1993) Science **262**:105-108.

Berget, S.M., Moore, C., and Sharp, P.A., (1977) Proc. Natl Acad. Sci. USA **74**:3171-3175.

Birnboim, H.C., and Doly, J. (1979) Nucleic Acids Res. **7**:1513-1523.

Blencowe, B.J., Sproat, B.S., Ryder, U., Barabino, S., and Lamond, A.I. (1989) Cell **59**:531-539.

Blumenthal, T. (1995) Trends Gen. **11**:132-136.

Bonneaud, N., Ozier-Kalogeropoulos, O., Li, G., Labouesse, M., Minvielle-Sebastia, L., and Lacroute, F. (1991) Yeast **7**:609-615.

Botstein, D., Falco, S.C., Stewart, S.E., Brennan, M., Scherer, S., Stinchcomb, D.T., Struhl, K., and Davis, R.W. (1979) Gene **8**:17-24.

Bradford, M.M. (1976) Anal. Biochem. **72**:248-254.

Brosi, R., Hauri, H.P., and Krämer, A. (1993a) J. Biol. Chem **268**:17640-17646.

Brosi, R., Gröning, K., Behrens, S.-E., Lührmann, R., and Krämer, A. (1993b) Science **262**:102-105.

Brown, J.D., and Beggs, J.D. (1992) EMBO J. **11**:3721-3729.

Brown, J.D. (1992) PhD Thesis, Edinburgh University.

Brown, J.W.S., Feix, G., and Frendeway, D. (1986) EMBO J. **5**:2749-2758

Burgess, S.M., and Guthrie, C. (1993) Cell **73**:1377-1391.

Burke, D., P. Gadaska, and L. Hartwell. (1989) Mol. Cell. Biol. **9**:1049-1059.

Carlson, M., Osmond, B.C., and Botstein, D. (1981) Genetics **98**:25-40.

Cavalier-Smith, T., (1985) Nature **315**:283-284.

Cech, T.R., *In* R. F. Gesteland and J. F. Atkins (ed.) The RNA World. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. (1993) p. 239-269.

Chanfreu, G., Legrain, P., Dujon, B., *et al.* (1994) Nucleic Acids Res. **22**:1981-1987.

Chang, T.-H., Clark, M.W., Lustig, A.J., Cusick, M., and Abelson, J. (1988) Mol. Cell. Biol. **8**:2379-2393

Chapon, C., and Legrain, P. (1992) EMBO J. **11**:3279-3288.

Cheng, S.-C., and Abelson, J. (1987) Genes Dev. **1**:1014-1027.

Cherest, H., Thomas, D., and Surdin-Kerjan, Y. (1990) Nucleic Acids Res. **18**:659.

Chien C.T., Bartel, P.L., Sternglanz, R., and Field, S. (1991) Proc. Natl Acad. Sci. **88**:9578-9582.

Chong, J.P.J., Mahbubani, H.M., Khoo, C.-Y., and Blow, J.J. (1995) Nature **375**:418-421.

Church, G.M., and Gilbert, W. (1984) Proc. Natl Acad. Sci. USA **81**:1991-1995.

Ciechanover, A. (1994) Cell **79**:13-21.

Clarke, P.R. (1995) Curr. Genetics **5**:40-42.

Cline, T.W. (1984) Genetics **107**:231-277.

Cooper, M., Johnston, L.H., and Beggs, J.D. (1995) EMBO J. **14**:2066-2075.

Couto, J.R., Tamm, J., Parker, R., and Guthrie, C. (1987) Genes Dev. **1**:445-455.

Cowan, P.M., and McGavin, S. (1955) Nature **176**:501-503.

Craig, E.A., Kramer, J., Shilling, J., Werber-Washburne, M., Holmes, S., Kosic-Smithers, J., and Nicolet, C.M. (1989) Mol. Cell. Biol. **9**:3000-3008.

Culbertson, M.R., and Winey, M. (1989) Yeast **5**:405-427.

Darnell, J.E., and Doolittle, W.F. (1986) Proc. Natl Acad. Sci. USA **83**:1271-1275

Datta, B., and Weiner, A.M. (1991) Nature **352**:821-824.

Davanloo, R., Rosenberg, A.H., Dunn, J.J., and Studier, F.W. (1984) Proc. Natl Acad. Sci. USA **81**:2035-2039.

Daveba, M.D., Post-Beittenmuller, M.A., and Wagner, J.A. (1986) Proc. Natl Acad. Sci. USA **83**:5854-5857.

Domdey, H., Apostol, B., Lin, R.-J., Newman, A.J., Brody, E., and Abelson, J. (1984) Cell **39**:611-621

Dowell, S.J., Romanowski, P., Diffley, J.F.X (1994) Science **265**:1243-1246.

Dreyfuss, G., Matunis, M.J., Pinol-Roma, S. *et al.* (1993) Ann. Rev. Biochem. **62**:289-321.

Dumas, L. B., J. P. Lussky, E. J. McFarland, and J. Shampay. (1982) Mol. Gen. Genet. **187**:42-46.

Eng, F.J., and Warner, J.R. (1991) Cell **65**:797-804.

Engebrecht, J.A., Voekel-Meiman, K., and Roeder, G.S. (1991) Cell **66**:1257-1268.

Espinoza, F. H., Ogas, J., Herskowitz, I., and Morgan, D.O. (1994) Science **266**:1388-1391.

Fabrizio, P., McPheeters, D.S., and Abelson, J. (1989) Genes Dev. **3**:2137-2150.

Fabrizio, P., and Abelson, J. (1990) Science **250**:404-409.

Feaver, W.J., Svejstrup, J.Q., Henry, N.L., and Kornberg, R.D. (1994) Cell **79**:1103-1109.

Feinberg, A.P., and Vogelstein, B. (1984) Anal. Biochem. **137**:266-267.

Fink, G.R. (1987) Cell **49**:5-6.

Fischer, U., Sumpter, V., Sekine, M., Satoh, T., and Lührmann, R. (1993) EMBO J. **12**:573-583.

Fleig, U.N., Pridmore, R.D., and Philippsen, P. (1986) **46**:237-245.

Flinn, E.M., (1994) PhD Thesis, Edinburgh University.

Fortner, D.M., Troy, R.G., and Brow, D.A. (1994) Genes Dev. **8**:221-233.

Fouser, L.A., and Friesen, J.D. (1986) Cell **45**:81-93.

Frank, D., Patterson, B., Guthrie, C. (1992) Mol. Cell. Biol. **12**:5197-5205.

Frendeway, D., and Keller, W. (1985) Cell **42**:355-367.

Friedman, D.B., Hollingsworth, N.M., Byers, B. (1994) Genetics **136**:449-464.

Fu, X.-D.,and Maniatis, T. (1990) Nature **343**:437-441.

Fu, X.-D. (1993) Nature **365**:82-85.

Fuller-Pace, F.V., and Lane, D.P. (1992) In F. Eckstein and D.M.J. Lilley, Nucleic Acids and Molecular Biology, Volume 6, Springer-Verlag, Berlin.

Gallwitz, D., and Sures, I. (1980) Proc. Natl Acad. Sci. USA **77**:2546-2550.

Garcia-Blanco, M.A., Jamison, S., and Sharp, P.A. (1989) Genes Dev. **5**:1224-1236.

Garcia-Blanco, M.A., Anderson, G.J., and Beggs, J.D. (1990) Proc. Natl Acad. Sci. USA **87**:3082-3086.

Ge, H., and Manley, J.L. (1990) Cell **62**:25-34.

Gerber, H.-P., Seipel, K., Georgiev, O., Höfferer, M., Hug, H., Rusconi, S., and Schaffner, W. (1994) Science **263**:808-811.

Gietz, D., St Jean, A., Woods, R.A., and Schiestl, R.H. (1992) Nucleic Acids Res. **20**:1425.

Gil, A., Sharp, P.A., Jamison, S.F., and Garcia-Blanco, M.A. (1991) Genes Dev. **5**:1224-1236.

Gilbert, W., Marchionni, M., and McKnight, G. (1986) Cell **46**:151-154

Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M., and Rochaix, J.-D. (1991) Cell **65**:135-143.

Green, M. (1986) Ann. Rev. Genet. **20**:671-708.

Green, M. (1991) Ann. Rev. Cell Biol. **7**:559-599.

Guialis, A., Moriatou, M., Patrinou-Georgoula, M., and Dangli, A. (1991) Nucleic Acids Res. **19**:287-296.

Guthrie, C. (1986) TIBS **11**:430-434.

Guthrie, C., and Patterson, B. (1988) Ann. Rev. Genet. **22**:387-419.

Guthrie, C., and Fink, G.R. (1991) Guide to Yeast Genetics and Molecular Biology. Vol.194, Methods in Enzymology. Academic Press Inc.

Hamilton, R., Watanabe, C.K., and de Boer, H.A. (1987) Nucleic Acids Res. **15**:3581-3593.

Hamm, J., Darzynkiewicz, E., Tahara, S.M., and Mattaj, I.W. (1990) Cell **62**:569-577.

Hanahan, D. (1983) J. Mol. Biol. **166**:557-580.

Hartwell, L.H. (1967) J. Bact. **93**:1662-1670.

Hartwell, L.H., McLaughlin, C.S., and Warner J.R. (1970) Mol. Gen. Genet. **109**:42-46.

Hausner, T.P., Giglio, L.M., and Weiner, A.M. (1990) Genes Dev. **4**:2146-2156.

He, F., Peltz, S.W., Donahue, J.L., Rosbash, M., Jacobson, A. (1993) Proc. Natl Acad. Sci. USA **90**:7034-7038.

Hermann, H., Fabrizio, P., Raker, V.A., Foulaki, K., Hornig, H., Brahms, H., and Lührmann, R. (1995) EMBO J. In press.

Hernandez, N., and Keller, W. (1983) Cell **35**:579-586.

Herskowitz, I. (1987) Nature **329**:219-222.

Hirling, H., Scheffner, M., Restle, T., and Stahl, H. (1989) Nature **339**:562-564.

Hodges, P. E., Jackson, S.P., Brown, J.D., and Beggs, J.D. (1995) Yeast **11**:337-342.

Hodges, P.E., and Beggs, J.D. (1994) Current Biology **4**:264-267.

Hopper, A.K., and Martin, N.C. (1992) *In* E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), The molecular and cellular biology of the yeast *Saccharomyces*, Volume 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. p. 99-141.

Hu, J., Xu, D., Schappert, K., Xu, Y., and Friesen, J. (1995) Mol. Cell. Biol. **15**:1274-1285.

Inoue, K., Hoshijima, K., Sakamoto, H., and Shimura, Y. (1990) Nature **344**:461-463.

Ito, H., Fukuda, Y., Murata, K., Kimura, A. (1983) J. Bact. **153**:163-168.

Jackson, A.L., Pahl, P.M.B., Harrison, K., Rosamond, J., and Sclafani, R.A. (1993) Mol. Cell. Biol. **13**:2899-2908.

Jackson, S.P. (1987) PhD Thesis, University of Edinburgh.

Jackson, S.P., Lossky, M., and Beggs, J.D. (1988) Mol. Cell. Biol. **8**:1067-1075.

Jacquier, A., and Rosbach, M. (1986) Proc. Natl Acad. Sci. USA **83**:5835-5839.

Jamieson, D.J., Rahe, B., Pringle, J., and Beggs, J.D. (1991) Nature **349**:715-717.

Johnson, M., and Davis, R.W. (1984) Mol. Cell. Biol. **4**:1440-1448.

Johnston, L.H., and Thomas, A.P. (1982a) Mol. Gen. Genet. **186**:439-444.

Johnston, L.H., and Thomas, A.P. (1982b) Mol. Gen. Genet. **186**:445-448.

Johnston, M. (1987) Micro. Rev. **51**:458-476.

Kandels-Lewis, S., and Séraphin, B. (1993) Science **262**:2035-2039.

Kiledjian, M., and Dreyfuss, G. (1992) EMBO J. **11**:2655-2664.

Kim, S.-H., Smith, J., Claude, A., and Lin, R.-J. (1992) EMBO J. **11**:2643-2654.

Kim, S.-H., and Lin, R.-J. (1993) Proc. Natl Acad. Sci. USA **90**:888-892.

King, D.S., and Beggs, J.D. (1990) Nucleic Acids Res. **18**:6559-6564.

Kitada, K., Johston, L.H., Sugino, T., and Sugino, A. (1992) Genetics **131**:21-29.

Klein, C., and Struhl, K. (1994) Mol. Cell. Biol. **14**:1920-1928.

Koch, C., and Nasmyth, K. (1994) Curr. Op. Cell Biol. **6**:451-459.

Kohtz, J.D., Jamison, J.F., Will, C.L., Zuo, P., Lührmann, R., Garcia-Blanco, M.A., and Manley, J.L. (1994) Nature **368**:119-124.

Konarska, M.M., and Sharp, P.A. (1986) Cell **46**:845-855.

Konforti, B.B., Koziolkiewicz, M.J., and Konarska, M.M. (1993) Cell **75**:863-873.

Krainer, A.R., Maniatis, T., Ruskin, B., and Green, M.R. (1984) Cell **36**:993-1005.

Krämer, A., Keller, W., Appel, B., and Lührmann, R. (1984) Cell **38**:299-307.

Kubota, Y., Mimura, S., Nishimoto, S.-I., Takisawa, H., and Nojima, H. (1995) Cell **81**:601-609.

Kulesza, H., Simpson, G.G., Waugh, R., Beggs, J.D., and Brown, J.W.S. (1993) FEBS Lett. **318**:4-6.

Laemmli, U.K. (1970) Nature **227**:680-685.

Lain, S., Reichmann, L.J,. and Garcia, J.A. (1990) Nucleic Acids Res. **18**:7003-7006.

Lamm, G.M., Blencowe, B.J., Sproat, B.S., Iribarren, A.M., Ryder, U., and Lührmann, R. (1991) Nucleic Acids Res. **19**:3193-3198.

Lamond, A. (1995) Pre-mRNA processing (ed.). Springer-Verlag, Berlin.

Langford, C., Nellen, W., Neissing, J., and Gallwitz, D. (1983) Proc. Natl Acad. Sci. USA **80**:1496-1500.

Larkin, J.C., Thompson, J.R., and Woolford, J.L. (1987) Mol. Cell. Biol. **7**:1764.

Lee, C.-G., and Hurwitz, G. (1993) J. Biol. Chem. **268**:16822-16830.

Legrain, P., Seraphin, B., and Rosbach, M. (1988) Mol. Cell. Biol. **8**:3755-3760.

Legrain, P., and Rosbach, M. (1989) Cell **57**:573-583.

Legrain, P., and Chapon, C. (1993) Science **262**:108-110.

Legrain, P., Chapon, C., and Galisson, F. (1993) Genes Dev. **7**:1390-1399.

Lehmeier, T., Raker, V., Hermann, H., and Lührmann, R. (1994) Proc. Natl Acad. Sci. USA **91**:12317-12321.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., and Steitz, J.A. (1980) Nature **283**:220-224.

Lesser, C.F., and Guthrie, C. (1993) **262**:1982-1988.

Liao, X.C., Tang, J., and Rosbach, M. (1993) Genes Dev. **7**:419-428.

Lin, R.-J., Newman, A.J., Cheng, S.-C., and Abelson, J. (1985): J. Biol. Chem. **260**:14780-14792.

Lin, R.-J., Lustig, A.J., and Abelson, J. (1987) Genes Dev. **1**:7-18.

Liu, H., J. Krizek, and A. Bretscher. (1992) Genetics **132**:665-673.

Lonberg, N., and Gilbert, W. (1985) Cell **40**:81-90.

Lossky, M., Anderson, G.J., Jackson, S.P., and Beggs, J.D. (1987) Cell **51**:1019-1026.

Lührmann, R., Kastner, B., and Bach, M. (1990) Biochim. Biophys. Acta **1087**:265-292.

Lustig, A.J., Lin, R.-J., and Abelson, J. (1986) Cell **47**:953-963.

Ma, J., and Ptashne, M. (1987) Cell **51**:113-119.

Ma, K., Inglis, J.D., Sharkey, A., Bickmore, W.A., Hill, R.E., Prosser, E.J., Speed, R.M., Thomson, E.J., Jobling, M., Taylor, K., Wolfe, J., Cooke, H.J., Hargreave, T.B., and Chandley, A.C. (1993) Cell **75**:1287-1295.

MacMillan, A.M., Query, C.C., Allerson, C.R., Chen, S., Verdine, G.L., and Sharp, P.A. (1994) Genes Dev. **8**:3008-3020.

Madhani, H.D., Bordonné, R., and Guthrie, C. (1990) Genes Dev. **4**:2264-2277.

Madhani, H.D., and Guthrie, C. (1992) Cell **71**:803-817.

Madhani, H.D., and Guthrie, C. (1994a) Annu. Rev. Genet. **28**:1-26

Madhani, H.D., and Guthrie, C. (1994b) Genes Dev. **8**:1071-1086.

Madine, M.A., Khoo, C.-Y., Mills, A.D., and Laskey, R.A. (1995) Nature **375**:421-424.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory.

Maschhoff and Padgett, R.A. (1993) Nucleic Acids Res. **21**:5456

Mattaj, I.W. (1986) Cell **40**:905-911.

Mattaj, I.W. (1988) In M. L. Birnstiel (ed.) Structure and function of the major and minor small nuclear ribonucleoprotein particles. Springer-Verlag, Berlin. P. 100-114.

Mayeda, A., and Krainer, A.R. (1992) Cell **68**:365-375.

McIntosh, E.M., Atkinson, T., Storms, R.K., and Smith, M. (1991) Mol. Cell. Biol. **11**:329-337.

McSwiggen, J.A., and Cech, T.R. (1989) Science **244**:679-683.

Measday, V., Moore, L., Ogas, J., Tyers, M., and Andrews, B. (1994) Science **266**:1391-1395.

Mermod, N., O'Neill, E.A., Kelly, T.J., and Tjian, R. (1989) Cell **58**:741-753.

Mermoud, J.E., Cohen, P., and Lamond, A.I. (1992) Nucleic Acids Res. **20**:5263-5269.

Mermoud, J.E., Cohen, P.T.W., and Lamond, A.I. (1994) EMBO J. **13**:5679-5688.

Messing, J., Gronenborn, B., Muller-Hill, B., and Hofschneider, P.H. (1977) Proc. Natl Acad. Sci. USA **74**:3642-3646.

Michaud, S., and Reed, R. (1991) Genes Dev. **5**:2534-2546.

Michaud, S., and Reed, R. (1993) Genes Dev. **7**:1008-1020.

Miller, A.M. (1984) EMBO J. **3**:1061-1065.

Mitchell, P.J., and Tjian, R. (1989) Science **245**:371-378.

Moore, M., and Sharp, P.A. (1992) Science **256**:992

Moore, M., and Sharp, P.A. (1993) Nature **365**:364

Moore. M.J., Query, C.C., and Sharp, P.A. *In* R. F. Gesteland and J. F. Atkins (ed.) The RNA world. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. (1993) p.303-357.

Moritz, M., Paulovich, A.G., Tsay, Y.-F., and Woolford, J.L. (1990) J. Cell. Biol. **111**:2261.

Morl, M., and Schmelzer, C., (1990) Cell **60**:629-636

Morran, (1990) PhD Thesis, Edinburgh University.

Mount, S.M. (1982) Nucleic Acids Res. **10**:459-472.

Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J.A. (1983) Cell **33**:509-518.

Mueller, M.W., Allmaier, M., Eskes, R., Schweyen, R.J. (1993) Nature **366**:174-176.

Nandabalan, K., Price, L.K., and Roeder, G.S. (1993) Cell **73**:407-415.

Nasmyth, K. (1993) Curr. Op. Cell Biol. **5**:166-179.

Navas, T.A., Zhou, Z., and Elledge, S.J. (1995) Cell **80**:29-39.

Nelissen, R.L.H., Will, C.L., van Venrooij, W.J., and Lührmann, R. (1994) EMBO J. **13**:4113-4125.

Nelson, K.K., and Green, M.R. (1990) Proc. Natl Acad. Sci. USA **87**:6253-6257.

Newman, A.J., Lin R.-J., Cheng, S.-C., and Abelson, J. (1985) Cell **42**:335-344.

Newman, A.J., and Norman, C. (1991) Cell **65**:115-123.

Newman, A.J., and Norman, C. (1992) Cell **68**:743-754.

Newman, A.J. (1994) Curr. Opinion Cell Biol. **6**:360-367.

Nomet, M., Sweetser, D., and Young, R.A. (1987) Cell **50**:909-915.

Ohya, Y., and Botstein, D. (1994) Science **263**:963-966; Genetics **138**:1041-1054.

Ozkaynak, E., Finley, D., Solomon, M.J., and Varshavsky, A. (1987) EMBO J. **6**:1429-1439.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984) Science **225**:898-903.

Parker, R.A., Silicano, P.G., and Guthrie, C. (1987) Cell **49**:229-239.

Parker, R., and Silicano, P.G. (1993) Nature **361**:660-662.

Parry, H.D., Scherly, D., and Mattaj, I.W. (1989) TIBS **14**:15-19.

Paterson, T., Beggs, J.D., Finnegan, D.L., and Lührmann, R. (1991) Nucleic Acids Res. **19**:5877-5882.

Patterson, B., and Guthrie, C. (1991) Cell **64**:181-197.

Patton, J.G., Mayer, A.M., Galceran, J., Tempst, P., and Nadal-Ginard, B. (1991) Genes Dev. **5**:1237-1251.

Patton, J.G., Porro, E.B., Galceran, J., Tempst, P., and Nadal-Ginard, B. (1993) Genes Dev. **7**:393-406.

Pearson, N.J., Thorburn, P.C., and Haber, J.E. (1982) Mol. Cell. Biol. **2**:571-577.

Peltz, S.W., Brown, A.H., and Jacobson, A. (1993) Genes Dev. **7**:1737-1754.

Pikielny, C.W., Teem, J.L., and Rosbach, M. (1983) Cell **34**:395-403.

Pikielny, C. W., Rymond, B.C., and Rosbach, M. (1986) Nature **324**:341-345.

Pines, J. (1994) Trends Biochem. Sci. **19**:143-145.

Pines, J. (1995) Nature **371**:742-743.

Pinto, A.L., and Steitz, J.A. (1989) Proc. Natl Acad. Sci. USA **86**:8742-8746.

Plessel, G., Fischer, U., and Lührmann, R. (1994) Mol. Cell. Biol. **14**:4160-4172.

Plumpton, M. (1993) PhD Thesis, Edinburgh University.

Plumpton, M., McGarvey, M., and Beggs, J.D. (1994) EMBO J. **13**:879-887.

Pon, L., and Schatz, G. (1991) *In* J.R. Broach, E.W. Jones, and J.R. Pringle. (ed.), The molecular and cellular biology of the yeast *Saccharomyces*, volume 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. (1991) p. 333-406.

Potashkin, J., Li, R., and Frendeway, D. (1989) EMBO J. **8**:551.

Query, C.C., Moore, M.J., and Sharp, P.A. (1994) Genes Dev. **8**:587-597.

Rao, H., and Stillman, B. (1995) Proc. Natl Acad. Sci. USA **92**:2224-2228.

Ray, B.K., Lawson, T.G., Thach, R.E., *et al.* (1985) J. Biol. Chem. **260**:7651-7658.

Reed, R., and Maniatis, T. (1985) Cell **41**:95-105.

Reed, R., and Maniatis, T. (1988) Genes and Dev. **2**:1268-1276.

Reed, R. (1989) Genes and Dev. **3**:2113-2123.

Reed, S.I. (1992) Ann. Rev. Cell Biol. **8**:529-561.

Reidel, N., Wise, J.A., Swerdlow, H., Mak, A., and Guthrie, C. (1986) Proc. Natl Acad. Sci. USA **83**:8097-8101.

Reidel. N., Wolin, S., and Guthrie, C. (1987) Science **235**:328-331.

Reifenberger, E., Freidel, K., and Ciriacy, M. (1995) Mol. Microbiol. **16**:157-167.

Ren, R., Mayer, B.J., Cicchetti, P., and Baltimore, D. (1993) Science **259**:1157-1161.

Richardson, J.S., and Richardson, D.C. (1988) Science **240**:1648-1652.

Robinson, P.A., Anderton, B.H., and Loving, T.L.F. (1988) J. Immunol. Meth. **108**:115-122.

Rodriguez, J.R., Pikielny, C.W., and Rosbach, M. (1984) Cell **39**:603-610.

Rodriguez-Medina, J.R., and Rymond, B.C. (1994) Mol. Gen. Genet. **243**:532-539.

Rogers, J., and Wall, R. (1980) Proc. Natl Acad. Sci. **77**:1875-1879.

Roy, J., Zheng, B., Rymond, B.C., and Woolford, J. (1995) Mol. Cell. Biol. **15**:445-455.

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., and Bowtell, D. (1993) Nature **363**:83-85.

Rozen, F., Pelletier, J., Trachsel, H., and Sonenburg, N. (1989) Mol. Cell. Biol. **9**:4061-4063.

Ruby, S.W., and Abelson, J. (1988) Science **242**:1028-1035.

Ruby, S.W., Chang, T.-H., and Abelson, J. (1993) Genes Dev. **7**:1909-1925.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984) Cell **38**:317-331.

Ruskin, B., and Green, M.R. (1985) Nature **317**:732-734.

Ruskin, B., Green, J.M., and Green, M.R. (1985) Cell **41**:833-844.

Ruskin, B., Pikielny, C.W., Rosbach, M., and Green, M.R. (1986) Proc. Natl Acad. Sci. USA **83**:2022-2026.

Rymond, B., Torrey, D., and Rosbach, M. (1987) Genes Dev. **1**:238-246.

Rymond, B.C., and Rosbach, M. (1992) *In* E. W. Jones, J. R. Pringle, and J. R. Broach (ed.), The molecular and cellular biology of the yeast *Saccharomyces*, Volume 2. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. p. 143-192.

Rymond, B.C. (1993) Proc. Natl Acad. Sci. USA **90**:848-852.

Sakai, A., Chibazakura, T., and Shimuzu, Y. (1992) Nucleic Acids Res. **20**:6227-6233.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning, A Laboratory Manual, second edition. Cold Spring Harbor Laboratory.

Santos, M.A., Garcia-Ramirez, J.J., Revuelta, J.L. (1995) J. Biol. Chem **270**:437-444.

Schatz, P. J., Pillus, L., Grisafi, P., Solomon, F., and Botstein, D. (1986a) Mol. Cell. Biol. **6**:3711-3721.

Schatz, P. J., Solomon, F., and Botstein, D. (1986b) Mol. Cell. Biol. **6**:3722-3733.

Schwer, B., and Guthrie, C. (1991) Nature **349**:494-499.

Schwer, B., and Guthrie, C. (1992) Mol. Cell. Biol. **12**:3540-3547.

Schwindinger, W.F., and Warner, J.R. (1987) J. Biol. Chem. **262**:5690.

Seipel, K., Georgiev, O., and Schaffner, W. (1992) EMBO J. **11**:4961-4968.

Sellem, C.H., Lecellier, G., and Belcour, L. (1993) Nature **366**:176-178.

Séraphin, B., Kretzner, L., and Rosbach, M. (1988) EMBO J. **7**:2553-2538.

Séraphin, B., and Rosbach, M. (1989) Cell **59**:349-358.

Séraphin, B., and Rosbach, M. (1990) Cell **63**:619-629.

Séraphin, B., Abovich, N., and Rosbach, M. (1991) Nucleic Acids Res. **19**:3857-3860.

Séraphin, B. (1995) EMBO J. **14**:2089-2098.

Seufert, W., and Jentsch, S. (1990) EMBO J. **9**:543-550.

Shannon, K.W., and Guthrie, C. (1991) Genes Dev. **5**:773-785.

Sharp, P.A. (1991) Science **254**:663

Sharp, P.A. (1994) Cell **77**:805-815.

Shea, J.E., Toyn J.H., and Johnston, L.H. (1994) Nucleic Acids Res. **22**:5555-5564.

Sheldon, M., and Reiberg, D. (1995) Current Biology **5**:43-46.

Shulman, R.W., and Warner, J.R. (1978) Mol. Gen. Genet. 161:221.

Sigler, P.B. (1988) Nature **333**:210-212.

Sikorski, R.S., and Heiter, P. (1989) Genetics **122**:19-27.

Silicano, P.G., Jones, M.H., and Guthrie, C. (1987) Science **237**:1484-1487.

Silicano, P.G., and Guthrie, C. (1988) Genes Dev. **2**:1258-1267.

Singh, R., and Reddy, R. (1989) Proc. Natl Acad. Sci. USA **86**:8280-9293.

Singh, R., Valcárcel, J., and Green, M.R. (1995) **268**:1173-1176.

Siomi, H., Matunis, M.J., Michael, W.M., and Dreyfuss, G. (1993) Nucleic Acids Res. **21**:1193-1198.

Smith, C.W.J., Porro, E.B., Patton, J.G., and Nadal-Ginard, B. (1989) Nature **342**:243-247.

Sontheimer, E.J., and Steitz, J.A. (1992) Mol. Cell. Biol. **12**:734-746.

Sontheimer, E.J., and Steitz, J.A. (1993) Science **262**:1989-1996.

Sosnowski, B.A., Belote, J.M., and McKeown, M. (1989) **58**:449-459.

Strauss, E. J., and Guthrie, C. (1991) Genes Dev. **5**:629-641.

Struhl, K. (1985) Proc. Natl Acad. Sci. **82**:8419-8423.

Suh, E.-R., and Waring, R.B. (1992) Nucleic Acids Res. **20**:6303-6309.

Swanson, M.S. (1995) In A. Lamond (ed.) Pre-mRNA processing. Springer-Verlag, Berlin. p. 17-29.

Teigelkamp, S., McGarvey, M., Plumpton, M., and Beggs, J.D. (1994) EMBO J. **13**:888-897.

Teigelkamp, S., Whittaker, E., and Beggs, J.D. (1995a) Nucleic Acids Res. **23**:320-326.

Teigelkamp, S., Newman, A.J., and Beggs, J.D. (1995b) EMBO J. **14**:2602-2612

Tollervey, D., and Mattaj, I.W. (1987) EMBO J. **6**:469-476.

Tollervey, D., Lehtonen, H., Carmo-Fonseca, M., and Hurt, E.C. (1991) EMBO J **10**:573-583.

Toyn, J.H., and Johnston, L.H. (1994) EMBO J. **13**:1103-1113.

Toyn, J.H., Toone W.M., Morgan, B.A., and Johnston, L.H. (1995) Trends Biol. Sci. **20**:70-73.

Tung, K., Norbeck, L.L., Nolan, S.L., Atkinson, N.S., and Hopper, A.K. (1992) Mol. Cell. Biol. **12**:2673-2680.

Tye, B.K. (1994) Trends Cell Biol. **4**:160-166.

Tyers, M., Tokiwa, G., Futcher, B. (1993) EMBO J. **12**:1955-1968.

Tzagoloff, A., and Myers, A.M. (1986) Ann. Rev. Biochem. **55**:249-285.

Umen, J. G., and Guthrie, C. (1995) Genes Dev., in press.

Utans, U., Behrens, S.-E., Kole, R., and Lührmann, R. (1992) Genes Dev. **6**:631-641.

Valcárcel, J., Singh, R., Zamore, P.D., and Green, M. (1993) **362**:171-175.

Valcárcel, J., Singh, R., and Green, M.R. (1995) *In* Lamond, A.I. (ed.) Pre-mRNA processing. Springer-Verlag, Berlin. p.97-112.

Vankan, P., McGuigan, C., and Mattaj, I.W. (1990) EMBO J. **9**:3397-3404.

Vijayraghavan, U., Parker, R., Tamm, J., Iimura, Y., Rossi, J., Abelson, J., and Guthrie, C. (1986) EMBO J. **5**:1683-1695.

Vijayraghavan, U., Company, M., and Abelson, J. (1989) Genes Dev. **3**:1206-1216.

Vijayraghavan, U., and Abelson, J. (1990) Mol. Cell. Biol. *10*:324-332.

Vilardell, J., and Warner, J.R. (1994) Genes Dev. **8**:211-220.

Wallace, J.C., and Edmonds, M. (1983) Proc. Natl Acad. Sci. USA **80**:950-954.

Wasserman, D.A., and Steitz, J.A. (1992) Science **257**:1918-1925.

Whittaker, E., and Beggs, J.D. (1991) Nucleic Acids Res. **19**:5483-5489.

Whittaker, E., Lossky, M., and Beggs, J.D. (1990) Proc. Natl Acad. Sci. USA **87**:2216-2219

Williams, K.A., and Deber, C.M. (1991) Biochem. **30**:8919-8923.

Williamson, M.P. (1994) Biochem. J. **297**:249-260.

Wise, J.A., Tollervey, D., Malony, D., Swerdlow, H,. Dunn, E.J., and Guthrie, C. (1983) Cell **35**:743-751.

Wittekind, M., Kolb, J.M., Dodd, J., Yamagishi, M., Memet, S., Buhler, J.-M., and Nomura, M. (1990) Mol. Cell. Biol. **10**:2049.

Wittowski, J.A., (1988) Trends in Biochem. Sci. **13**:110-113.

Woolford, J.L., and Warner, J.R. (1991) *In* J.R. Broach, E.W. Jones, and J.R. Pringle. (ed.), The molecular and cellular biology of the yeast *Saccharomyces*, volume 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. p. 587-626.

Woppmann, A., Will, C.L., Kornstädt, U., Zuo, P., and Manley, J.L. (1993) Nucleic Acids Res. **21**:2815-2822.

Wu, J., and Manley, J. (1989) Genes Dev. **3**:1553-1561.

Wu, J.A., and Manley, J.L. (1991) Nature **352**:818-821.

Wu, J.Y., and Maniatis, T. (1993) Cell **75**:1061-1070.

Wyatt, J.R., Sontheimer, E.J., and Steitz, J.A. (1992) Genes Dev. **6**:2542-2553.

Yean, S.L., and Lin, R.-J. (1991) Mol. Cell. Biol. **11**:5571-5577.

Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W., and Schreiber, S.L. (1994) Cell **76**:933-945.

Zahler, A.M., Lane, W.S., Stolk, J.A., and Roth, M.B. (1992) Genes Dev. **6**:837-847.

Zahler, A.M., Neugebauer, K.M., Lane, W.S., Roth, M.B. (1993) Science **260**:219-222.

Zamore, P.D., and Green, M.R. (1989) Proc. Natl Acad. Sci. USA **86**:9243-9247.

Zamore, P.D., Patton, J.G., and Green, M.R. (1992) Nature **355**:609-614.

Zeitlin, S., and Efstradiatis, A. (1984) Cell **39**:589-602.

Zhuang, Y., and Weiner, A.M. (1986) Cell **46**:287-835.

Zhuang, Y., and Weimer, A.M. (1989) Genes Dev. **3**:1545-1552.

Zillman, M., Gorowski, M.M., and Phizicky, E.M. (1991) Mol. Cell. Biol. **11**:5410-5416.

Acknowledgements

I am deeply grateful to my supervisor Jean, for her time, support and enthusiasm over the past four years, and for her patience with my at times tangential approach to research. The rapid reading and re-reading of thesis chapters I particularly appreciated. The members of the Beggs lab, past and present, have made life in the lab (and about the town) thoroughly enjoyable, and have put up with me admirably. Thanks to the excellent travelling companions Stefan and Susan (who tolerated my map reading, but only just), to the eccentrics Liz and Rob for adding spice to life, to the french contingent Valérie and Sandrine for much wonderful cooking, to Mary for not being able to put weight behind a punch, to Fiona (and Rocky the cat), and to any others with whom I have enjoyed Edinburgh licensing laws over the years: Nik, Jim, Margaret, Mike, Sarah, the babies Andrew and Ian, and Caroline. A mention also for the tee-totallers (or nearly so): Michelle, for being such a whipper-snapper, and Big Ian, for enduring Chapter 2.

I am indebted to all those who have contributed ideas and constructive criticism of my work; special thanks go to Nik and to Peter and the other Postdocs in the lab.

I dedicate this thesis to my parents, for their unending support and dedication, and for forgiving me for not heeding their advice (not to enter Science).

Finally, an enormous thank you to Heather, the best travelling companion of all, without whom I would have surely gone insane and given up long ago.

This work was supported by a Studentship from the Cancer Research Campaign.

Notes:

(1): "JDY" plasmids were constructed by J.D.Brown.

(2): As described in Brown & Beggs (1992), manipulation of *PRP8* involved directed mutagenesis to introduce an *Nhe*I site immediately after the start codon. This entailed two nucleotide substitutions, causing one amino acid change at the position 5 (proline to alanine). In the table, pY8000 and pY8500 bear the wild-type chromosomal allele (no *NheI* site), and in all other (later) plasmids *PRP8* contains the *Nhe*I site. For simplicity, and as this mutagenesis has not been detected to affect the activity of the protein (Brown and Beggs, 1992), both versions of *PRP8* (with and without *Nhe*I) are regarded as wild-type in this thesis.

(3): Deletions (Δ) are defined by amino acid number.

Appendix A

Table A.1: Plasmids Used In This Study.

(A): Vectors

Name	Derived from	Features
pRS316	pBluescript	*CEN, ARS, URA3* Multiple Cloning Site
pBM125	YCp50	*CEN, ARS, URA3*
YEp24	-	2μ, *URA3*

(B): *PRP8* alleles

Name[1]	Features	Derrived from	Allele[2,3]	Copies of LP_nG	Promoter
pY8000	2μ, *URA3, PRP8*	YEp24	wild-type	4	*PRP8*
pY8500	*CEN, ARS, URA3, PRP8*	YCp50	wild-type	4	*PRP8*
pJDY6	2μ, *URA3, PRP8*	YEp24	wild-type	4	*PRP8*
pAJY12	*CEN, ARS, URA3, PRP8*	pRS316	wild-type	4	*PRP8*
pJDY10	*CEN, ARS, URA3, GAL1UAS:PRP8*	pRS316	wild-type	4	hybrid *GAL1UAS:PRP8*
pJDY13	*CEN, ARS, HIS3, GAL1:PRP8*	pRS316	wild-type	4	*GAL1* (glucose repressed, weak expression)
pBM-PRP8	*CEN, ARS, URA3, GAL1:PRP8*	pBM125	wild-type	4	*GAL1-10* (overexpression)
pAJY11	*CEN, ARS, URA3, GAL1:PRP8*	pRS316	wild-type	4	*GAL1* (overexpression)
pAJY13	*CEN, ARS, URA3, prp8ΔP2*	pRS316	Δ29 to 79	2	*PRP8*
pAJY1	*CEN, ARS, URA3, GAL1:prp8ΔP2*	YEp24	Δ29 to 79	2	*GAL1* (overexpression)
pJDY7	2μ, *URA3, prp8ΔN*	YEp24	Δ4 to 79	none	*PRP8*
pAJY9	*CEN, ARS, URA3, GAL1:prp8ΔN*	pBM125	Δ4 to 79	none	*GAL1* (overexpression)

Appendix B: N-terminal sequence of *dbf3-1* suppressor cDNAs.

Number	N-terminal Sequence (length in nucleotides).
1	GACCCACGCG TCCGTAAAAA TGGCCTCCTT ACCTCACCCA AAGATTGTCA AGAAGCACAC CAAGAAGTTC AAGCGTCATC ACTCTGACCG TTACCACAGA GTTGCTGAAA ACTGGAGAAA GCAAAAGGGT ATTGACTCTG TTTGTTAGAA GAAGATTCAG AGGTAACATC TCTCAACCAA AGATCGTTAC G (191).
2	CGCTTTTAAA CATCCTGCAA ACAGTCTAAT AAATACGTAT AATAAGTTTG AAATGGGAGG CATTTCGTGA AAAGAAAGCT GAATACTTTG CTAAATTAAG AGAATACTTG GAAGAATACA AGT (123).
3	CGTCCGCCCA CGCGTCCGGT AACAATGGCT ATCTCCAAGA ATTTACCAAT TTTGAAGAAC CACTTGAGAA AGCACTGGCA AGAACGTGTC AAGGTTCACT TTGACCAAGC NGGTAAAAAG GTTTCTAGAC GTAATGCTAG AGCTGCCAGA GCCGCCAAGA TTGCTCCAAG ACCATTGGAT TTGTTGAGAC CTGTTGTCAG AGCTCCAACT GTTAAGTAAC AACAGAAAGG TTAGAGCTGG TAGAGGTTTA CCTTGGCTGA AG (262).
6	CCCACGCGTC CGCTAAAAAT ATCACCGGAA TCTGTCTTTG GGACATTAGG TGAACTGATT AAAGTCAATG ATAAGTACAA GACATGTGCT GAAGTAATTG GAGGTAACTC CTCTTCCACT TGTCGTGGAT ACCGTAGAAA CGGCCACTTT GATTATGAAC GAACTATATC GTATTGAAGG GTGGGAGAGT GACGTTTATT CCC (203).
7	GCCGTCGCTG CTTTAGAAGC TGCTAGAGAC AGCAAGTCCC CAATCATTTT GCAAACCTCT AACGGTGGTG CTGCTTACTT CGCTGGTAAG GGTATCTCTA ACGAAGGTCA AAATGCTTCC ATCAAGGGTG CTATTGCCGC TGCCCACTAC ATCAGA (156).

Number	N-terminal Sequence (length in nucleotides).
8	CCACGCGTCC GAAATGGGTG TTGAACAATC TTAAAGAGAA AGCCGGTGTC ATCGTTGGTG AAGATGTCCA CAACTTATTC ACTTCGCTAA GGAACACAAG TTCG (104).
10	GCTCCTAACC TGTATGATTG TCTTTGCCTG TTTCTATATT TTTTTTTTTG CTACCACTTG (60).
12	CCCACGCGTC CGAATATCCA ATCCTTTATT TTTTTGCAGA AATGAAATTC CAACTGCCGT TACTACGTTG ATTAGTTCTG GTGCCATCGT CTCTGCGGGA CCACACGTTG GATGTTCACC AAGAAGATGC CCACCACCAT AAGAGGGCCG TTGCGTACAA ATACGTTTACGAAACTGTTG TTGTCG (186).
13	CCCACGCGTC CGCCCACGGT CCGGATAATG TCATTTGATA GACCGGAATA TACAGTGCGC CTGTTTTACA AGGAGAATCT CCTAACGACG ATGATAATAC TGAAATCATA AAGTCCTTTA AGAATTTCAT TTGGAGTTCA GACTTGACTC GCAATTTATT TACAGAGATC AGTTAAGGAA CAACATCCT (189).
14	GATGAATTGA GATCCATGGT TAGAAAGTGG CAAACTTTGA TCGAAGCTAT ACGTTGACTG TTAAGACTTC TGATGATTAC GTTTTGAG (88).
16	GGCAAACCAT TAGTTCATTC GAAAGAACGT ATTGTCGAGA ATTATCATTC ACTATATCAG AAAATTGACT ACTACG (76).
17	GTTATCGGTC ATGTCGATTC TGGTAAGTCT ACCACTACCG GTCATTTGAT TTACAAGTGT GGTGGTATTG ACAAGAGAAC CATCGAAAAG TTCGAAAAGG AAGCCGCTGA A (111).

Number	N-terminal Sequence (length in nucleotides).
18	GCTGAAAAGA CCGGTAAGCG TCAAGTTTTA ATAGACCATC CTCCAAGGTC ATTAACAAGT TTTTGCAAGT TATGCAAAAG CACGGTTACA TTGGTGAATT TGAATACATC GATGACCACA GATCTGGTAA CA (132).
19	CCTGCCGGCC CTGTCAGGTC TTTGGAAGGG GCAGCCACTA TCCCAGTGTC ACTGACCACA GCAGGCTTGG TGTTGACCTA TAACTTGGGC TTGAACCTGA AGTGGGAGCC ATCAACCCCA CAAAGAAAC (129).
20	TTGAATCTAA CGCATGGGAT GGTAGAGACG CCATTGTAGT TTGCGGTGAT ATTGCCATCT ACGATAAGGG TGCCGCAAGA CCAACCGGTG GTGCCGGTAC TGTTGCTATG TGGATCGGTC CTGATGCTCC AATTG (135).
22	CGACTACGCT GCTTACATGT TCAAGTACGA CTCCACTCC (39).
23	GTTGCTTTGA ACGACCCATT CATCACCAAC GACTACGCTG CTTACATCAG TACGACTCCA CTCACGGTAG ATTCGCTGGT GAAGTTTCCA CGATGACAAG CACATCATTG TCGA (114).
24	TGTTGGTCAT GATGCAACTC TCCCGTCGCA TTGACATGGA GGACCCAACC ATCATCATGT ACATTAGAAT TTTATACTGT TCTTGGATCG GTATCTCTTG GATCATCTAC CAAATCGGCC ACAAAGAGAA TTGTTGCTAA AAACGACATG ACTACCATGA AGTACCGTCG AACCTGGTAA TGC (183).
25	GACCCACGCG TCCGAAGCAG CGATCAACAA ATTCCAGACT TTGAACAAGC TGATAAAAAT AGAGAAAACC A (71).
26	CCCACGCGTC CGATTTTAAT TGAACAAAAT GGCTGAAGGT GTTTCCAAGG TGCTATCGGT ATCGATTTAG GTACAACCTA CTCTTGTGTT

Number	N-terminal Sequence (length in nucleotides).
	GCTACTTACG AATCCTCCGT TGAAATTATG GCCAACGAAC AAGGTAACAG AGTCACCCCA TCTTTCGTTG CTTTCACTCC AGAAGAAAGA TTCATTGGTG ATGCTGCCAA GAACCAA (207).
27	TCGTGCCTTC AAAAGGAAAA GGTTGAAGAA CAAGAACAAC AACAACAACA AATCATCAAG ATTCGATCCA CTTTGACCTC CACCAAGGTT AAGCAATTGG AAAACG (106).
30	TTTATTACCG CTTAAACTGA TTATTTCAAA AAAAAAAAGT AATATGTATT AATCAACAAC AGAAATGAAA TCTTAGCTGC TTACTTATTA TTGGTTCAAG GTGGTAAC (108).
31	TGGGTTTGAT GAAGACCAAG TTAAAGGTG CCAACCAAGG TTATTCTTTA CTAATGCGT (59).
32	GTCCGGCAAT CATACCCCTC ATGATCGCTA TTGGATTATT TATGGCACAA CCCTTCGGTG GTCATTCGCT CCTTCAAACA AGTTTTTCTA CGATATTAAC AGATGATTTG AAAACTCAAA AGGTAGCTAT AAAGGAATTA GAAAGGAAAA TTGCTGAGAT GGACCCGGGG GCTTCAGTTG TGACTAAATG CTTGAAGGAG GG (202).
33	TCATTGGTGA AGATGTCACA ACTTATTCAC TTACGCTAAG GAACACAAGT TCGCTATTCC AGCTATTAAC GTCACCTCTT CTTCTACTGC CGTCGCTGCT TTACAAGCTG CTAGAGACAG CAAGTCCCCA ATCATTTTGC AAACCTCTAA CGGTGGTGCT GCTTACTTCG CTCCTAAGGG TATCTCTAAC GAAGGTCAAA ATGCTTCC (208).
34	CACGGGGAAC TTGTGCAATG TTTGCAGGAA ATAAGAGAAG TTTAGCAAGG AGGCCAATAT CCAATCACAA GATCCTACCT CTAAGCTTGG CAAGGTACTA GACTGTACGT (160).

Number	N-terminal Sequence (length in nucleotides).
36	CGCCATTGCT GCCCCACGTT CAAGTGAACG ATTACAGACG ACAAGGAACA ACCAGAAGAG TCCAGGGGGC CAACAGCAAC A (81).
37	ACAACAATTA CCGTTTCTCC TTCACACAAC ATGGCTTCTA CTTCGAACAC GTTCCCTCCA AGCCAAAGCA ATTCTTCCAA CAACCTTCCA ACTTCTAGAC ATGCATCCAT TGTGGAGA (118).
41	CGGCTGGAAA GGTCTAATTA ATCACCCTGA TCTTAACAAC ACTTTCAACA TCAACAAGGG TTTGCAATCC GCTAGACAAT TGTTGTCAAC TTGACAATAT (100).
42	TCTTTACCAG CTACTTTTGA TTTCAGTCCA GAGGATGCCC AACTTTTCTT ATACGTTTTC AATGCTAGAC CAGATGGTGT T (131).
43	TTTCCGTTAA GGAAATCAGA AGAGGTAACG TCTGTGGTGA CGCTAAGAAC GATCCACCCA AAGGGTTG (68).
44	GAGCTGAAGC TTTGAACATT TTCTGGTGAA TTCTTC (36).
45	GATTAAAATT TCTCTTCCAT TCTTTGACGT TTTAAATCAT ACCTCGTATT CTTTACGCAA AAGAAAGATA CCTGTGAAAT ATATCTGAAT ATGGTCCGAT TA (102).
47	TTTTATTCTC AATGTCAACC CATTTGATTT GTTAGGTAAC GAGTTCAGAT GACGAAGACG TTTTGGCCAC CAAAGGAAAT CGTAAGAGCA ACACTTCCTC CAAGAAGGCT G (111).

Number	N-terminal Sequence (length in nucleotides).
48	CCCGAAACAA ACAAAATGGT TAGAGTTGCT ATTAACGGTT TCGGTAGAAT CGGTAGATTG GTCATGAGAA TTGCAAACTC TAGACCAAAC GTCGAAGTTG TTGCTTTGAA CGACCCATTC ATCACCAACG ACTAC (135).
50	CACCGTGGAG CTCCAATTCG (20).
51	AAAAACAGTA CTTCACTAAA TTTACACACA AAACAAAATG ATCATAATTG CTATTAACGG TTTCGGTACA ATCGGTAGAT T (81).
52	TCCAATGGGT TTCTCACCAT TGGATGGTTC TAAATCAACC AACGAAAATG TATCTGCTTC CACTTCTACT GCCAAACCAA TGGTTGGCCA ATTG (94).
55	AAAAATGGCC TCCTTACCTC ACCCAAAGAT TGTCAAGAAG CACACCAAGA AGTTCAGCGT CATCACTCTG ACCGTTACCA CAGAGTTGCT GAAAACTGGA GA (102).
57	CGTCTAATAT CGGCGCATAT TTATGAAATC TAAAAGGCAA AGAATCTTAA TACATAAAAT CCCAGACGGG GGATGGAATT CAAAGTGTCA CCATTGACAA AGATTATTTC TCTGTCAGGA TTCCTCGC (128).
X	CCCACGCGTC CGCCCACGCG TCCGCTTACA ACTGCAAACA ACAATGAGAA GTTATTAGTA TTAATGTCGG TCAAGCTGGT TGTCAGATTG GTAATGCCTG TTGGGAATTA TATTCCCGTG AGCACGGTAT TAAGCCGGAT GGACATCTAG AAGATGGCCC TTTCAAAGCC GAAGGAGGGA GAA (183).