Title
Genetic and biochemical analysis of interactions involving Prp8 and RNA at the catalytic core of the spliceosome

Author
Collins, Catherine A.

Publication Date
2000

Peer reviewed|Thesis/dissertation

Genetic and Biochemical Analysis of Interactions Involving Prp8 and RNA at the Catalytic Core of the Spliceosome

by

Catherine A. Collins

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA SAN FRANCISCO

Date University Librarian

Degree Conferred: .....

For my family:

Donald, Vicki, Edward, Amy, and Michael Collins,

and Clara Pfoff

ACKNOWLEDGMENTS

Chapter 1 is a reprint from:

Collins, C. A., and Guthrie C. (1999). Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes and Development* 13: 1970-82.

The Epilogue is a reprint from:

Collins, C. A., and Guthrie C. (2000). The question remains: is the spliceosome a ribozyme? *Nature Structural Biology*, vol. 7: 850-4.

I am tremendously indebted to my advisor Christine Guthrie and members of the Guthrie lab for all that they have taught and shared with me.

Genetic and Biochemical Analysis of Interactions Involving Prp8 and RNA at the Catalytic Core of the Spliceosome

Catherine A. Collins

ABSTRACT

A long-standing question is whether the pre-mRNA splicing reaction is catalyzed by RNA. A candidate protein component of the spliceosome's active site, or catalytic core, is Prp8. This is the most highly conserved (and largest) spliceosomal protein and forms numerous crosslinks to spliceosomal RNA, including highly conserved consensus sequences immediately adjacent to both the 5' and 3' splice sites. Genetic properties of specific alleles of Prp8 have functional implications for the interactions of Prp8 with RNA in the catalytic core. Remarkably, one class of *prp8* alleles can simultaneously suppress a discrete subset of mutations in both the 5' splice site and the 3' splice site consensus sequences, and in a specific residue within the essential ACAGAG motif of U6 snRNA. The results are best explained by a model in which Prp8 influences a tertiary interaction involving all of the affected RNA residues. This interaction provides new constraints for modeling the catalytic core and provides further support for the hypothesis that the catalytic core is primarily an RNA structure, which may be stabilized or chaperoned during its formation by protein.

To determine whether the RNA residues influenced by Prp8 indeed form direct interactions with each other, I screened for allele-specific genetic interactions among residues in the 5' splice site, 3' splice site, and U6. While systematic analysis rules out any standard base pairing or base triple interactions, one striking example of genetic suppression was observed involving a specific combination of mutations in the 5' splice site, 3' splice site, and U6.

In a parallel biochemical approach I have developed a crosslinking assay to probe interactions of the 3' splice site consensus sequence with RNA. Results from this assay should provide physical constraints for modeling the catalytic core at the second step of splicing. Preliminary results are consistent with previous genetic results, which suggest that residues in U2 snRNA also participate in the tertiary interaction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Page

Appendix

INTRODUCTION

A critical step in gene expression is the removal of intervening sequences, called introns, from messenger RNA. This process, called pre-mRNA splicing, is catalyzed by the spliceosome, a large RNA-protein complex consisting of five small nuclear RNAs (snRNAs), the intron-containing pre-mRNA substrate, and about 60 proteins (Moore et al., 1993). The splicing reaction consists of two independent phosphoryl transfer reactions, or steps (Figure 1). For the first, the 5' phosphate of the intron (the 5' splice site), is attacked by a 2' hydroxyl group within the intron (the so-called branchpoint). The liberated 3' exon then attacks the 3' phosphate of the intron (the 3' splice site) for the second phosphoryl transfer step. The final result of the splicing reaction is ligation of the two exons (to generate mRNA) and excision of the intron in a branched ('lariat') form.

The catalytic core of the spliceosome

Ever since the discovery that RNA is capable of performing catalysis (Cech and Bass, 1986), an intriguing hypothesis has been that the spliceosome is an RNA enzyme. By the time I began my thesis work, many compelling similarities had been noted between the spliceosome and a class of RNA enzymes called Group II introns (Michel and Ferat, 1995; Pyle, 1996; Weiner, 1993). These introns, found interrupting some prokaryotic and mitochondrial genes, can catalyze their own splicing (with the help of metal ions and in some cases, protein cofactors). The splicing mechanism for both the spliceosome and Group II introns proceeds via two transesterification steps that follow the same stereochemical course (Moore and Sharp, 1993; Padgett et al., 1994). Also, both reactions

utilize a 'branchpoint' 2' hydroxyl from a conserved adenosine within the intron as a nucleophile for the first step. Other shared features include regions of secondary structure and the function of a stem loop in positioning exons for ligation (Michel and Ferat, 1995; Pyle, 1996; Weiner, 1993). Since my work began, it has been shown that both the spliceosome and Group II introns are metalloenzymes (Gordon et al., 2000; Sontheimer et al., 1999; Sontheimer et al., 1997), and similarities have been noted in the structural and sequence context of RNA residues implicated in binding metals (Chanfreau and Jacquier, 1994; Yean, 2000). The high degree of similarity in the reactions suggest that the spliceosome and Group II introns have evolved from a common ancestor.

While the pre-mRNA splicing reaction may have once been catalyzed by RNA, the question remains, is it still catalyzed by RNA? How far has the present day machinery diverged from an RNA-based catalytic strategy? Do any of the spliceosomal proteins play critical catalytic functions?

Understanding the mechanism of catalysis requires understanding the components and structure of the so-called catalytic core, or active site, of this complicated machine. Previous work had identified a number of critical RNA interactions within the spliceosome between the snRNAs with one another and with the pre-mRNA substrate, that could possibly play critical roles in catalysis (Figure 2, top). U2 and U6 snRNAs interact with the intron near two of the three sites of chemistry—the 5' splice site (5'SS) and the branchpoint (BP)—*via* the 5'SS helix and BP helix, respectively. The adjoining U2/U6 helices I, II and III could help juxtapose these two reactive sites for the first

phosphoryl-transfer step (blue arrow). The highly conserved loop of U5 snRNA had been hypothesized (and since then, implicated) in positioning the exons for ligation during the second phosphoryl transfer step (indicated by c in Figure 2). Additional long-range interactions suggested the existence of a more compact RNA tertiary structure. These include an interaction between the first and last Gs of the intron (interaction a), and an invariant U6 residue near the 5'SS helix with a bulged U2 residue in helix I (interaction b).

While the catalytic core is clearly comprised of RNA, it might also contain protein. One particular spliceosomal protein has stood out from the others as a candidate component of the catalytic core. Numerous crosslinking studies indicate that Prp8 physically contacts RNA residues near the sites of catalysis: the 5' splice site, 3' splice site, and branchpoint (highlighted in light blue in Figure 2) (Chiara et al., 1996; Chiara et al., 1997; MacMillan et al., 1994; Teigelkamp et al., 1995; Umen and Guthrie, 1995; Wyatt et al., 1992). Consistent with a fundamental role in splicing, Prp8 is the largest (280 kDa) and most highly conserved splicing protein identified, with 61% identity between yeast and worm homologues, throughout its entire sequence of ~2400 amino acids (Hodges et al., 1995). The work presented in this thesis verifies that Prp8 performs important functions in conjunction with RNA at the catalytic core.

The dynamic spliceosome assembly process

Importantly, the catalytic core of the spliceosome is very dynamic. It does not exist in the absence of the intron-containing substrate. (Indeed, it is argued in the following chapters that intron residues are important components of the active site). The spliceosome must assemble anew upon each intron to be spliced from subparticles called snRNPs, which contain individual snRNAs and numerous tightly associated proteins.

In vitro characterization of the spliceosome assembly process had identified many steps (Figure 3). The U1 and U2 snRNPs appear to assemble onto the intron first. Critical for their assembly are base pairing interactions between U1 snRNA with the 5' splice site consensus sequence, and U2 snRNA with the branchpoint. Multiple sub-steps in the formation of the pre-spliceosome are thought to ensure that introns are recognized accurately before becoming committed to the splicing pathway. This 'pre-spliceosome' is then joined by a complex of the U4, U5 and U6 snRNPs called the 'triple snRNP'. As a component of the U5 snRNP and the triple snRNP, this is the point at which Prp8 joins the spliceosome, as well as U6 snRNA, which is a key catalytic component.

Before catalysis can proceed, a dramatic conformational rearrangement among the snRNP components takes place. In order to form the interactions of the catalytic core, other interactions within the snRNPs must be disrupted. For example the base pairing interactions that form U2/U6 helix I and the intramolecular U6 helix are excluded by the extensive base pairing of U4 with U6 snRNA within the triple snRNP, and a U2

intramolecular helix within the U2 snRNP (hatched in Figure 2). The base pairing of U1 snRNA with the 5' splice site must be disrupted in order for U6 snRNA to interact to interact with the 5' splice site.

These rearrangements may function to regulate the formation of the catalytic core, or perhaps as a pathway for chaperoning the folding of the catalytic core. The timing and coordination of disruption and formation of each helix with respect to others is likely to be important, yet it has been difficult to pinpoint the relative timing of individual interactions. A number of spliceosomal factors, including a family of ATPases (so-called DEAD box proteins) which are related in sequence to molecules that catalyze RNA unwinding, have been implicated in mediating these RNA rearrangements (Staley and Guthrie, 1998).

Evidence for conformational rearrangements (and requirement for ATP hydrolysis) exists not only for catalytic activation, but for nearly every stage in figure 3. Even the fully assembled spliceosome is not static, since several biochemical assays have detected evidence for at least one conformational change between the two steps site (Schwer and Guthrie, 1992; Umen and Guthrie, 1995). The exact nature of any of the individual rearrangements requires a more detailed understanding of the structure of preceding and following intermediates.

The dynamic nature of the spliceosome provides great challenges to biochemical study. It is difficult to purify any one spliceosomal sub-state from multiple other sub-

states that form in the course of a splicing reaction. In contrast, there is no barrier to probing the functional consequences of introducing specific mutations. Described in this thesis, a genetic approach has uncovered new interactions in the catalytic core.

A genetic approach to identifying functions for Prp8

As an integral component of the U5 snRNP, U4/U5/U6 triple snRNP, and the spliceosome, Prp8 is privy to many of the RNA rearrangements involved in the assembly of the catalytic core. My initial goal was to identify one or more specific roles for this highly conserved protein in one or more sub-steps of the splicing reaction. Remarkably, the sequence of Prp8 contains no identifiable protein motifs, precluding any testable hypotheses for its biochemical activities. However, the crosslinking interactions with spliceosomal RNAs suggested that Prp8 might function via interactions with a specific RNA sequence or structure.

A specific hypothesis was that Prp8 may recognize the 5' and/or 3' splice site consensus sequences. Previously, only snRNAs had been implicated in such a critical function. U1, U2, and U6 snRNA function recognize the 5' splice site and branchpoint consensus sequences through base pairing interactions (reviewed in Burge et al., 1998; Moore et al., 1993). Interestingly, no snRNA base pairing interaction has been identified to interact with the short but essential 3' splice site consensus UAG. Although Prp8 crosslinks to this sequence (Chiara et al., 1996; Chiara et al., 1997), it also crosslinks to residues in the 5' and 3' exons, whose sequence identity is not conserved (Teigelkamp et al., 1995). Thus it was an open question whether the crosslinking interactions of Prp8

reflected functions in recognizing specific sequences, rather than mere physical proximity to the RNA residues probed.

A former graduate student in the lab, Jim Umen, found functional evidence to implicate Prp8 in recognizing the 3' splice site, by identifying specific alleles of *prp8* that could suppress the splicing defect of mutations in the 3' splice site consensus UAG (Umen and Guthrie, 1996). He also found another specific class of *prp8* alleles that impaired the recognition of the poly-pyrimidine tract that proceeds the 3' splice site UAG in most yeast introns. Interestingly, the two classes of alleles mapped to two distinct regions of the Prp8, suggesting the existence of separable functional domains

Jim Umen's work encouraged me to continue a genetic approach to identifying functions for Prp8. As an integral component of the U5 snRNP (which contains about 20 proteins) and the U4/U6/U5 triple snRNP (Brown and Beggs, 1992), the activities of Prp8 in the spliceosome could not be investigated through traditional biochemical depletion and complementation experiments. However, the specific alleles identified by Jim could presumably assemble into spliceosomes, and revealed a later requirement for Prp8 in the second step.

Because numerous crosslinking studies suggest that Prp8 interacts with the 5' splice site, my thesis begins with an extension of Jim Umen's genetic analysis to functionally probe these interactions. I focussed upon the hypothesis that Prp8 recognizes the highly conserved GU dinucleotide adjacent to the 5' splice site, because these residues were observed to form a particularly efficient UV crosslink to Prp8

(Reyes et al., 1996). In the process of addressing this hypothesis, I found some interesting answers to the following questions.

1) Does Prp8 play a *functional* role in recognizing the 5' splice site GU? This would be provided by the identification of *prp8* alleles that permit splicing of introns which contain mutations in the 5' splice site GU. Indeed, I found a bounty of *prp8* alleles that could suppress mutations in position 2, the U, of the 5' splice site GU.

2) What region of Prp8 is functionally implicated in 5' splice site recognition? I screened the entire coding region of Prp8 and predicted that, as for the phenotypes identified by Jim Umen, 5' splice site suppressor mutations would map to a distinct region of Prp8. However, I found suppressor mutations in four different regions of Prp8. One overlaps in location with Jim's 3' splice site suppressor alleles, and another overlaps in location with the pyrimidine-tract recognition alleles. Thus, the simple expectation that one region of Prp8 functions in 5' splice site recognition, and another in 3' splice site recognition, was incorrect and has been replaced by a more complex and interesting view.

3) Does the 5' splice site suppression phenotype actually reflect a physical interaction of Prp8 with the 5' splice site GU? We addressed this question in a collaborative effort with Magda Konarska's lab at Rockefeller University. Using a protease fingerprinting approach, they mapped the 5' splice site GU crosslink to a region of 10 amino acids on Prp8 (Reyes et al., 1999). I helped them conduct more saturating screens of mutations in the limited region of Prp8 near the GU crosslink for a splice site suppression phenotype. They identified alleles that map close to the GU crosslink which show very similar phenotypes to my alleles identified from screening the entire (~2400

amino acids) coding region of Prp8. Thus, the same suppression phenotype can arise from mutations close to the site of crossinking, and elsewhere, suggesting that suppression may indeed occur via a direct interaction between Prp8 and the GU dinucleotide at the 5' splice site.

4) Is 5' splice site suppression phenotypically distinct from the other specific *prp8* phenotypes identified by Jim Umen? Jim had previously reported that the 3' splice site suppressor alleles did not suppress mutations in the 5' splice site, but the mutations tested were in 5' splice site positions 1 and 5; position 2 was not tested. When I assayed the 3' splice site suppressor alleles for effects upon mutations in 5' splice site position 2, I was surprised to find that they behaved very similarly to my 5' splice site suppressor alleles. Conversely, my 5' splice site suppressor alleles were able to suppress mutations in the 3' splice site. Thus it appeared that the two 'splice site suppression' phenotypes are actually the same, reflecting one function for Prp8 at both splice sites.

The most attractive explanation for effects by one Prp8 mutation upon distinct intron sites was the existence of a tertiary interaction between the 5' and 3' splice sites, which would allow Prp8 to influence both sites simultaneously. Consistent with this idea, the *prp8* alleles appear to suppress both 5' and 3' splice site mutations at the same step (the second step) of the splicing reaction. The tertiary interaction we proposed, when considered with other observations, predicted that U6 snRNA would also be involved. Indeed, I found that the *prp8* alleles also suppress mutations in a specific U6 residue that would be expected to participate in the hypothesized tertiary interaction.

The above genetic results are described in Chapter 1, a re-print from a publication in *Genes and Development*. They implied, to our initial surprise, that Prp8 does not recognize an individual RNA sequence. Rather, Prp8 appears to influence an RNA structure. In retrospect, this is a believable function for a protein at the catalytic core of the spliceosome. This study set the stage for some difficult but important questions for further understanding the structure and mechanism of the spliceosome's catalytic core: what is the nature of the proposed tertiary interaction that is influenced by Prp8? Is it primarily an RNA interaction or an intimate RNA-protein interaction? How does Prp8 interact with RNA? Chapters 2 and 3 describe two parallel approaches (using genetics and biochemistry) towards obtaining further evidence for the tertiary interaction that is implied by the genetic interactions of Prp8.

In Chapter 2, I carried out genetic studies to look for a specific tertiary RNA structure. That is, if a standard base pair or triple forms between 5' and 3' intron RNA residues, then specific combinations of mutations may be able to suppress each other. This approach has the advantage that such an observation could allow one to make specific and testable predictions about the nature of interactions in the spliceosome. I screened through many combinations of RNA mutations and found one striking three-way genetic interaction involving residues in the 5' splice site, 3' splice site, and U6 snRNA, supporting the hypothesis of a tertiary interaction. However, a specific direct physical interaction must still be proven. A manuscript describing this work is currently being prepared.

Chapter 3 describes the development of a photo-crosslinking assay for identifying RNA contacts to the 3' splice site UAG. The proposed tertiary interaction suggests that the 3'splice site should crosslink to the U6 snRNA and/or the 5' splice site. After developing successful crosslinking conditions, I identified two 3' splice site crosslinks which form with a kinetics indicative of the second step of splicing. One crosslink appears to be to U2. Neither appears to be to U6. Further work is needed to identify whether the other is to the 5'splice site.

The Epilogue (a reprint from *Nature Structural Biology*) reviews our current understanding of the protein and RNA contributions to spliceosome catalysis. Prp8 is clearly an important protein component of the catalytic core, and we cannot rule out the possibility that it participates directly in the catalytic reaction. However, the evidence that Prp8 influences a tertiary RNA interaction provides even more support for the idea that the catalytic core is fundamentally an RNA structure that is stabilized or chaperoned by Prp8.

More open-ended approaches towards understanding interactions at the catalytic core lie ahead for future intrepid explorers of the function and structure of Prp8. An initial start and some future directions towards identifying Prp8's RNA binding domain are described in the Appendix. Also, highlighted in the Epilogue, recent data from the Manley lab (and from Tommaso Villa in our lab) suggest that exciting new approaches towards dissecting the structural and functional contributions of individual spliceosome components may be possible through reconstitution experiments.

Given the results of Chapters 1 and 2, I have often worried that the RNA structure bound by Prp8 might be refractory to study. I hope that my thesis work has helped prepare us for the next stage. There is now quite heartening hope, and terrific people who are carrying that hope to the bench, that the interactions of Prp8 with RNA may indeed someday be studied in structural detail.

Figure 1: The pre-mRNA splicing reaction consists of two phosphoryl-transfer steps.

In the first step, the 5' intron phosphate, (the 5' splice site) is attacked by a 2' hydroxyl specified within the intron (the branch point). In the second step, the 3' intron phosphate (the 3' splice site) is attacked by the 3' hydroxyl of the cleaved 5' exon. The final result of the splicing reaction is ligation of the two exons (to generate mRNA) and excision of the intron in a branched ('lariat') form. The reaction catalyzed by self-splicing Group II introns can be described with the same illustration (because the two reactions are so similar in mechanism).

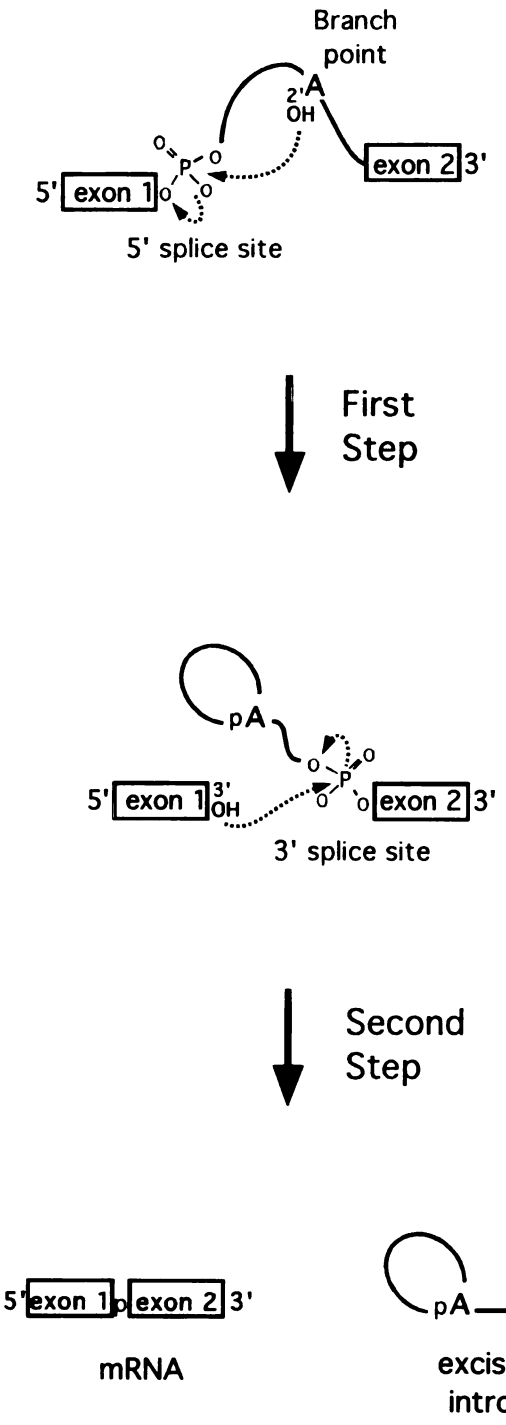

Branch
point
2'
OH
A
exon 2 3'
5' exon 1
5' splice site
First
Step
pA
5' exon 1 3'
OH
exon 2 3'
3' splice site
Second
Step
5' exon 1 exon 2 3'
mRNA
pA
3'
OH
excised
intron

Figure 2: Model of RNA and Prp8 interactions in the catalytic core, and mutually exclusive interactions of U6 and U2 snRNAs in snRNPs.

Grey lines denote Watson-Crick basepairing interactions. Exons are denoted by rectangles, and the intron substrate is in black. Essential U6 sequences, the ACAGAG motif, and the AGC are indicated (Madhani et al., 1990). The location of residues in U6 and U2 snRNA that participate in mutually exclusive base pairing interactions in the spliceosome, and the snRNPs, are hatched. Green dotted lines indicate tertiary interactions (a), (b), and (c).

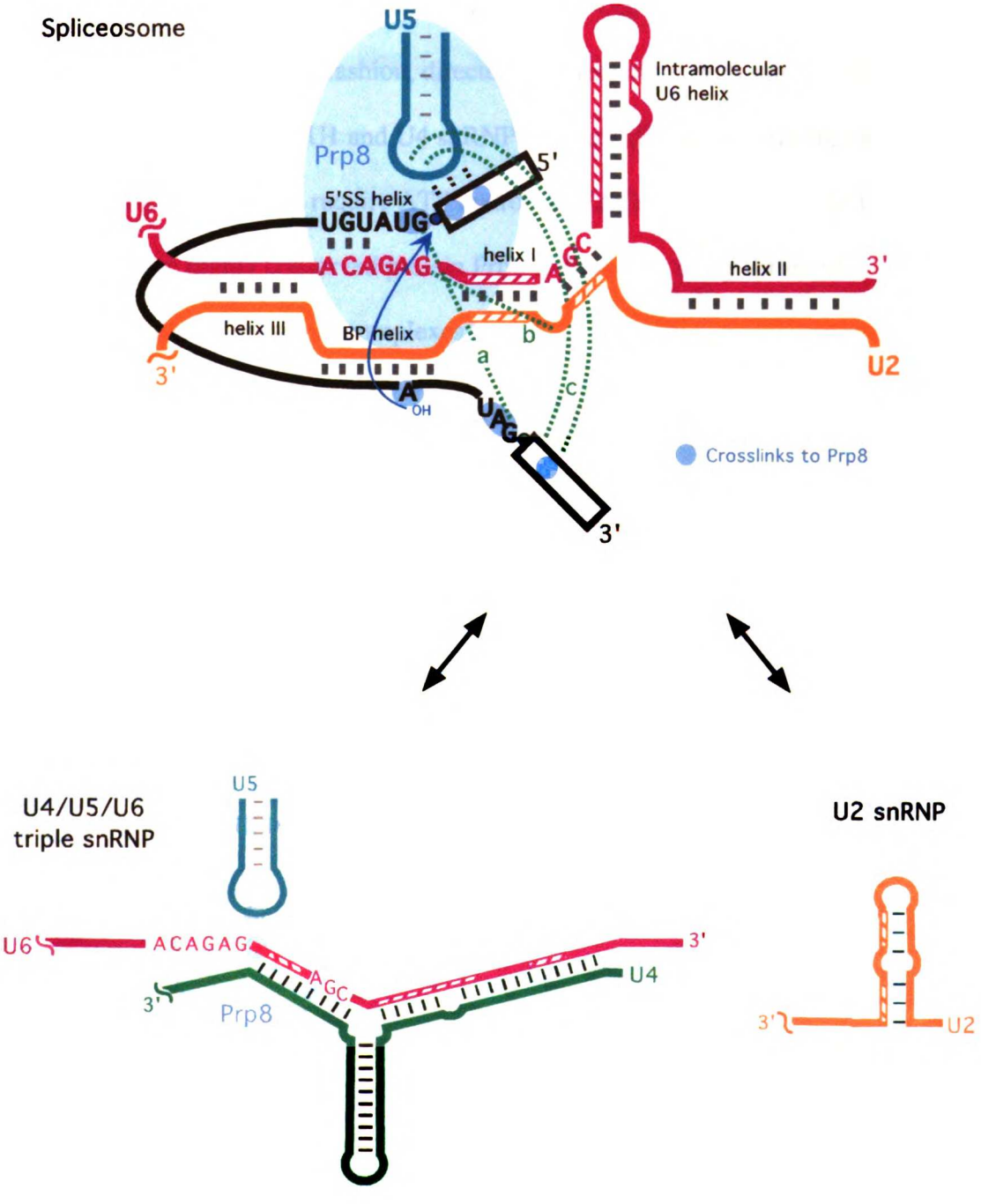
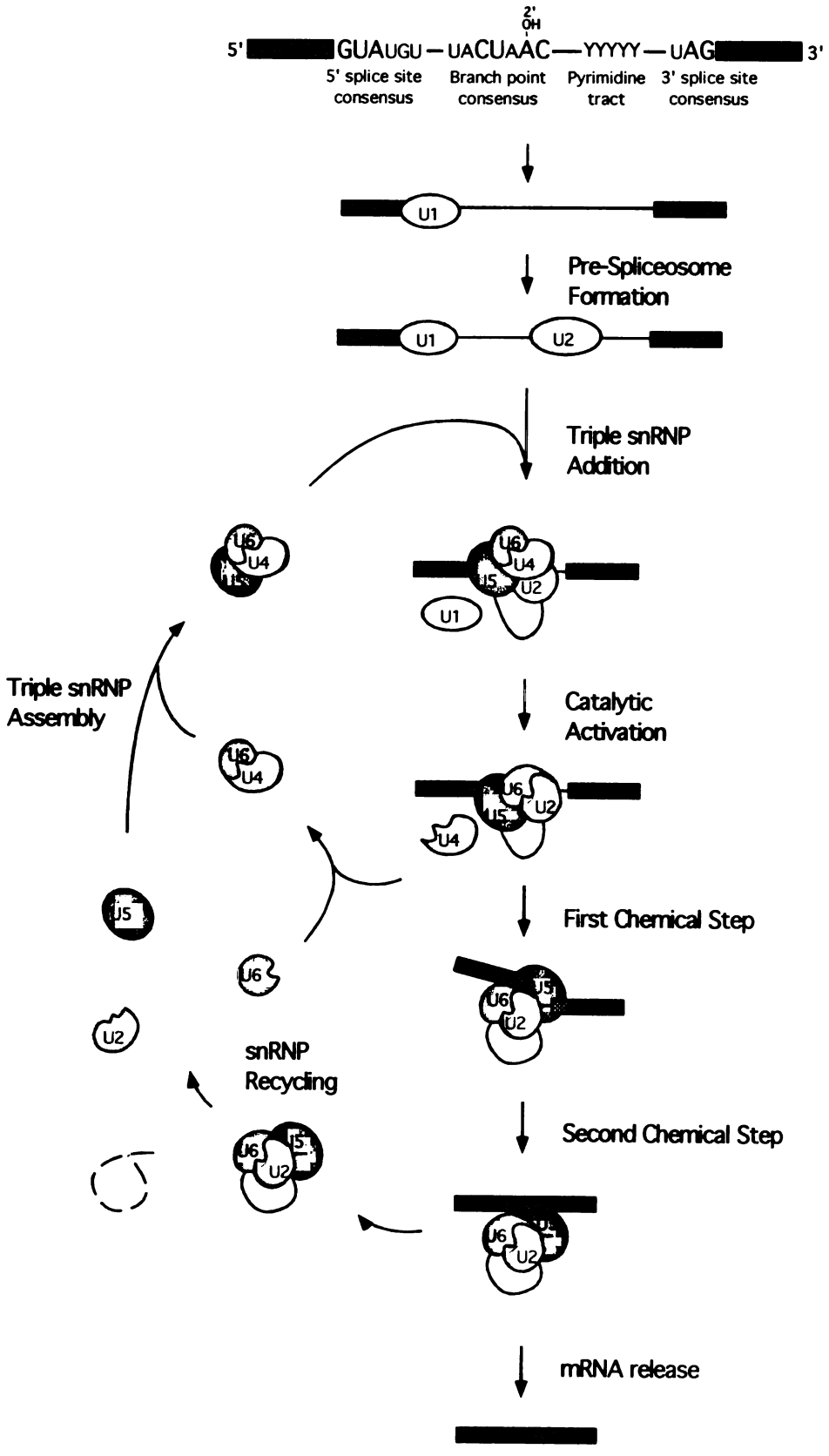

Spliceosome
U5
Prp8
5'SS helix
UGUAUG
ACAGAG
U6
5'
3'
helix I
helix II
helix III
BP helix
Intramolecular
U6 helix
U2
A
OH
UAG
a
b
c
Crosslinks to Prp8
U4/U5/U6
triple snRNP
U5
U6
ACAGAG
AGC
U4
Prp8
U2 snRNP
U2

Figure 3: Stages of spliceosome assembly.

Small nuclear ribonuclear protein particles, snRNPs, (U1, U2, U4, U5, U6) assemble onto the intron substrate in a step-wise fashion, directed by interactions with intron consensus sequences and one another. The U1 and U4 snRNPs become destabilized during or before catalytic activation. Many steps require ATP hydrolysis. As a component of the U5 snRNP, (in blue), the highly conserved protein Prp8 is present in many different spliceosomal and sub-spliceosomal complexes.


2'
OH
5'
GUAUGU — UACUAAC — YYYYY — UAG
3'
5' splice site consensus
Branch point consensus
Pyrimidine tract
3' splice site consensus
U1
Pre-Spliceosome Formation
U1
U2
Triple snRNP Addition
U6
U4
U5
U2
U1
Catalytic Activation
U6
U4
U5
U2
First Chemical Step
U6
U5
U2
Second Chemical Step
U6
U5
U2
mRNA release
Triple snRNP Assembly
U6
U4
U5
U6
U2
snRNP Recycling
U6
U5
U2

CHAPTER 1

Allele-specific genetic interactions suggest a function for Prp8 at the catalytic core of the spliceosome

ABSTRACT

The highly conserved spliceosomal protein Prp8 is known to crosslink the critical sequences at both the 5' (GU) and 3' (UAG) ends of the intron. We have identified *prp8* mutants with the remarkable property of suppressing exon ligation defects due to mutations in position 2 of the 5' GU, and all positions of the 3' UAG. The *prp8* mutants also suppress mutations in position A51 of the critical ACAGAG motif in U6 snRNA, which has previously been observed to crosslink position 2 of the 5' GU. Other mutations in the 5' splice site, branchpoint, and neighboring residues of the U6 ACAGAG motif, are not suppressed. Notably, the suppressed residues are specifically conserved from yeast to man, and from U2- to U12-dependent spliceosomes. We propose that Prp8 participates in a previously unrecognized tertiary interaction between U6 snRNA and both the 5' and 3' ends of the intron. This model suggests a mechanism for positioning the 3' splice site for catalysis, and assigns a fundamental role for Prp8 in pre-mRNA splicing.

INTRODUCTION

The removal of introns from eukaryotic genes to generate mRNA is catalyzed by the spliceosome, a complex RNA-protein machine composed of small nuclear RNAs (snRNAs), and at least sixty proteins. The pre-mRNA splicing reaction consists of two sequential transesterification steps: the first step results in cleavage of the 5' splice site and formation of a branched ("lariat") intermediate; the second step results in ligation of the 5' and 3' exons. The existence of self-splicing group II introns, which utilize a similar two-step transesterification mechanism, has led to the hypothesis that pre-mRNA splicing is catalyzed by RNA (Cech, 1986; Sharp, 1985). Indeed, mutational and crosslinking analyses have revealed a network of RNA-RNA interactions, which have been proposed to form the catalytic core of the spliceosome (Newman, 1994; Nilsen, 1994).

Whether any of the spliceosomal proteins make direct structural or chemical contributions to the active site remains unknown. Numerous crosslinking studies have placed one spliceosomal protein, Prp8, at or near the catalytic core. As summarized in Figure 1, Prp8 has been observed to crosslink to RNA residues within each critical sequence component of the intron: the consensus sequences that define the 5' splice site (5'SS), the branch-point (BP), and the 3' splice site (3'SS) (Chiara et al., 1996; Chiara et al., 1997; MacMillan et al., 1994; Reyes et al., 1996; Teigelkamp et al., 1995; Umen and Guthrie, 1995). Prp8 also forms extensive crosslinks to residues in both the 5' and 3' exons (Chiara et al., 1996; Teigelkamp et al., 1995; Wyatt et al., 1992) and to U5 snRNA (Dix et al., 1998), which interacts with the 5' and 3' exons (Newman and Norman, 1992; Sontheimer and Steitz, 1993). Prp8 is distinguished by its unusually high evolutionary conservation: 62%

sequence identity from yeast to man over the entirety of its ~2500 residues (Hodges et al., 1995; Lucke et al., 1997; Luo et al., 1999). Prp8 is thus hypothesized to play a fundamental role at the active site. Unfortunately, the sequence of this large protein provides no clues to its domain organization or its potential biochemical activities.

We have taken a genetic approach, in the budding yeast *S. cerevisiae,* to identify functional correlates to Prp8's crosslinking interactions with RNA. This approach exploits the *ACT1-CUP1* splicing reporter system, through which the splicing of mutant introns can be measured *in vivo* by growth on copper (Lesser and Guthrie, 1993). This system has previously been used to provide functional support for Prp8's observed interactions with the 3' splice site (Umen and Guthrie, 1996). Through screens of randomly mutagenized *prp8* alleles, mutants were identified (*prp8-121* through *-125*) which suppress the splicing defect of *ACT1-CUP1* reporters containing mutations in the 3'SS UAG consensus sequence. This finding suggested a role for Prp8 in recognition of this sequence during the exon ligation step. A separate class of *prp8* mutants suggested a role for Prp8 in recognition of the pyrimidine tract that precedes most 3' splice sites. It was known from *in vivo* studies of a construct containing two 3' splice sites in competition that the 3' splice site following a pyrimidine tract is preferentially used (Patterson and Guthrie, 1991); the *prp8-101* through *–107* alleles were isolated in a screen for loss of this preference. Interestingly, *prp8* mutations that give rise to the two classes of phenotypes, loss of pyrimidine tract preference (Pyr), and 3'SS UAG suppression, map to distinct regions of the protein and are functionally non-overlapping. This suggests that different domains of *PRP8* can be mutated to uncover separable functions.

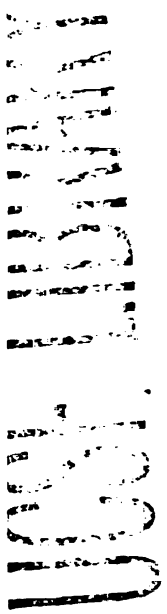
Our current study was motivated by the identification of a very strong UV crosslink between the human Prp8 homologue, p220, and the invariant GU di-nucleotide at the 5' end of the intron (5'SS GU) (Reyes et al., 1996). Our goals were twofold. First, we sought functional verification of a role for Prp8 in binding the 5'SS GU, by looking for suppression of mutations at the 5' splice site. Second, we wanted to know whether the protein location of such 5' splice site suppressors would implicate a new functional *PRP8* domain.

We have identified alleles of *prp8* which suppress the splicing of mutations in position 2 of the 5'SS GU. Suppression arises from viable mutations in at least four regions of the protein, including the two regions previously implicated in 3'SS UAG and Pyr phenotypes. Surprisingly, these alleles concomitantly suppress mutations in the 3'SS UAG. Moreover, we found that all of the originally identified 3'SS UAG suppressor alleles also suppress 5'SS position 2 mutations. This newly defined *prp8* phenotype of 5'SS and 3'SS suppression does not extend to other mutations in the 5'SS or in the branchpoint consensus sequences. However, mutations in A51 of the critical U6 ACAGAG motif, which crosslinks to 5'SS position 2 (Kim and Abelson, 1996; Sontheimer and Steitz, 1993), are also suppressed. All of the mutations suppressed confer a strong defect to exon ligation, which is partially rescued by the *prp8* alleles. We suggest that suppression of the distinct subset of RNA mutations could occur through a loosening of an active site structure comprised of a previously unrecognized tertiary interaction. The proposed interaction, whose components are conserved from yeast to man, and from U2 to U12 dependent spliceosomes, could serve to position the 3' splice site for catalysis. Our findings are consistent with the view that the spliceosomal catalytic core is

fundamentally comprised of RNA and that the highly conserved protein, Prp8, binds a critical RNA structure at the spliceosomal active site.

RESULTS

Screen for prp8 alleles that affect fidelity for the 5'SS GU



To determine if Prp8 functions at the 5' SS GU, we screened for *prp8* alleles that relax or change the requirement for these residues in splicing. Using *ACT1-CUP1* reporters (Lesser and Guthrie, 1993) that contained mutations in 5'SS position 2, we mutagenized *PRP8* and screened for alleles that confer growth on higher concentrations of copper, reflecting an increased efficiency of splicing. *PRP8* mutagenesis was conducted with Mn^{2+}-enhanced PCR (Leung et al., 1989), and a gap repair strategy was used to introduce plasmid-borne mutagenized *prp8* into strains, as depicted in Figure 2A (Muhlrad et al., 1992). To reduce the frequency of null mutations, and to facilitate the mapping of mutations, we separately mutagenized each of four equal fragments (A-D) of the *PRP8* coding region (Umen and Guthrie, 1996). For each fragment, approximately 4,000 mutant transformants were screened. The strain used was deleted for the chromosomal *PRP8* and wild type *PRP8* was supplied on a counter-selectable *URA3* marked plasmid. After loss of this plasmid by growth on 5-FOA, approximately 60% of the *prp8* mutant transformants were unable to support viability. We screened all transformants before 5-FOA passage for dominant suppression and the surviving 40% after 5-FOA passage for recessive suppression. The screens yielded *prp8* suppressors of

two 5'SS position 2 mutations, U2A and U2G, from the B, C, and D mutagenized fragments. These new *prp8* alleles confer a modest but reproducible increase in copper resistance: a two- to four- fold increase for U2A and barely a two- fold increase for U2G (Figure 2B). All of the *prp8* alleles identified exhibit dominant suppression and are haploviable.

Mapping the GU suppressor alleles

While Prp8's domain structure is not known, the distinct location of mutations that confer the previously characterized Pyr and 3'SS UAG phenotypes suggested the existence of separable functional domains. Thus, we were curious whether 5'SS position 2 suppression arises from mutations in similar regions of Prp8, or whether another distinct functional domain could be defined. Of particular interest was whether any of these suppressor mutations localize to the site of 5'SS GU crosslinking on Prp8 (Reyes et al., 1999).

The 5' SS position 2 suppressors lie in four regions of Prp8, indicated in black bars in Figure 3. One cluster of suppressor mutations, identified from screens with the D mutagenized fragment, lie 100 residues from the site of 5'SS GU crosslinking. The clustering of mutations is unlikely to be due to bias in the mutagenesis procedure, because many alleles contained additional mutations outside these regions that did not contribute to the suppression phenotype (data not shown). Interestingly, the cluster of 5'SS position 2 suppressor alleles from the D mutagenized fragment overlaps in location

with the 3'SS Pyr alleles, and another cluster from the C mutagenized fragment overlaps in location with the 3'SS UAG suppressor alleles.

Allele specificity of 5'SS and 3'SS suppression

To address whether suppression of mutations in the 5' splice site consensus sequence and suppression of mutations in the 3' splice site consensus sequence are separable or overlapping phenotypes, we analyzed the allele specificity of suppression. Each *prp8* allele was tested for effects upon a number of different mutations in the 5'SS and 3'SS sequences. A mutation in the intron residue that becomes the nucleophile for 5'SS cleavage, within the so-called branchpoint (BP) consensus sequence, was also tested for suppression by *prp8* alleles. Copper growth conferred upon strains bearing a series of mutant *ACT1-CUP1* reporters are shown for representative *prp8* alleles in Figure 4A. The phenotypes for all alleles are summarized in Figure 4B.

Interestingly, we found that all of the *prp8* alleles that suppress 5'SS position 2 also suppress at least one 3'SS UAG mutation. Conversely, the previously identified 3'SS UAG suppressor alleles all suppress the 5'SS U2*A* mutation. Importantly, suppression by these *prp8* alleles does not extend to all mutations in the intron, because the 5'SS mutations G1A and G5A, as well the BP mutation A259C, are not suppressed. Thus, the effect of the *prp8* alleles upon splicing is specific to a discrete, although unanticipated, subset of mutations.

While some of the 5'SS position 2 suppressor alleles contain multiple mutations, the phenotypes at the two sites do not arise from separate mutations. For some of the

5'SS suppressors (Table 1) and all of the 3'SS suppressors, (Umen and Guthrie, 1996), a single mutation that confers suppression has been identified. We conclude that 5'SS position 2 and 3'SS UAG suppression are actually the same phenotype, reflecting a single function for Prp8 at both sites. This phenotype, which we now simply call "splice site suppression", is distinct from the previously characterized Pyr phenotype, since *prp8* alleles that are defective for pyrimidine tract recognition do not suppress any intron mutation (Figure 4, and data not shown). Likewise, none of the splice site suppressor alleles affect recognition of the pyrimidine tract ((Umen and Guthrie, 1996), and data not shown).

That the same mutations on *prp8* give rise to phenotypes at both the 5'SS and 3'SS suggests that the same part(s) of the protein functions at both sites. Whether the location of the suppressor mutations reflect the location of protein-RNA contacts is not known. Two observations, in addition to the striking specificity, suggest that suppression is not simply due to indirect effects upon splice site residues. One is that the degree of suppression for different 3'SS UAG mutations is distinct for each *prp8* allele (Fig. 4A, and (Umen and Guthrie, 1996)). Another is that, in the companion study, *prp8* mutations in residues very close to the site of 5'SS GU crosslinking give rise to the same "splice site suppression" phenotype (Siatecka et al., 1999).

The dual suppression by the same *prp8* mutations also suggests that position 2 of the 5'SS interacts with the 3'SS UAG. Consistent with this notion, strong evidence for an interaction between 5'SS position 1 and the last residue of the 3' splice site has been reported (Chanfreau et al., 1994; Deirdre et al., 1995; Parker and Siliciano, 1993). However, there are no available data implicating 5'SS position 2 in an interaction with any residues

in the 3'SS UAG. In fact, evidence for an interaction with the 3'SS penultimate residue has been sought, but not found (Ruis et al., 1994), conceivably due to additional constraints imposed by Prp8.

prp8 alleles suppress the exon ligation defect of both 5'SS position 2 and 3'SS UAG mutations

The suggested interaction might explain the longstanding observation that mutations in 5'SS position 2 confer a strong defect to exon ligation, the result of the second catalytic step of splicing, (Aebi et al., 1986; Aebi et al., 1987; Fouser and Friesen, 1986), as do mutations in the 3'SS UAG (Chanfreau et al., 1994; Fouser and Friesen, 1987; Parker and Siliciano, 1993; Reed and Maniatis, 1985; Ruskin and Green, 1985; Vijayraghavan et al., 1986). Since the *prp8* splice site suppressor alleles rescue the exon ligation defect of 3'SS UAG mutations (Umen and Guthrie, 1996), they might concomitantly rescue the exon ligation defect of 5'SS position 2 mutations. To test this hypothesis, we assayed the *in vivo* efficiency of the exon ligation step for the U2A and U2G mutant reporters by primer extension analysis (Figure 5). The *prp8* splice site suppressor alleles allowed an increase in the steady-state levels of mRNA. From phosphorimage analysis, the ratio of spliced (mature) RNA to lariat-intermediate RNA was calculated in order to estimate the efficiency of the exon ligation step (Fouser and Friesen, 1986; Pikielny and Rosbash, 1985). For all alleles except *D-144* (see below), this efficiency was increased for both U2A and U2G splicing from two- to twenty- fold (Figure 5B and data not shown). The increase in the efficiency of exon ligation could reflect a direct effect upon the second catalytic step

of splicing, or on any substep in the spliceosomal rearrangements that must occur between the first and second catalytic steps.

5'SS position 2 mutations also confer a modest defect to steps preceding and possibly including 5' splice site cleavage (Aebi et al., 1987; Fouser and Friesen, 1986; Konforti and Konarska, 1994; Siatecka et al., 1999). Notably, the estimated overall efficiency of U2A and U2G splicing by the *prp8* alleles (Figure 5C) correlates, as expected, with the degree of copper resistance observed for these reporters (Figure 4A), but does not exactly correlate with the estimated efficiency of exon ligation (5B). The differences between the estimates in figures 5C and 5B could be due to additional effects upon other steps of splicing, such as those preceding or including 5' splice site cleavage. In a concurrent study, *in vitro* analysis suggests that *prp8* splice site suppressor alleles do indeed confer a modest increase in the efficiency of the 5' splice site cleavage step, by enhancing the association of Prp8-containing U4/5/6 triple snRNPs with the mutant transcript during the assembly of spliceosomes (Siatecka et al., 1999).

In our *in vivo* analysis, one splice site suppressor allele, *prp8-144*, does not appear to suppress the exon ligation step. Consistent with only a weak effect on exon ligation, this allele exhibits barely detectable suppression of 3'SS UAG mutations (Figure 4A). Because this *prp8* allele does increase the overall efficiency of 5'SS U2G splicing (Figure 5C), we suspect that the distinct location of *prp8-144*'s mutations allows for suppression at steps prior to but not including exon ligation.

prp8 alleles suppress a mutation in U6 snRNA

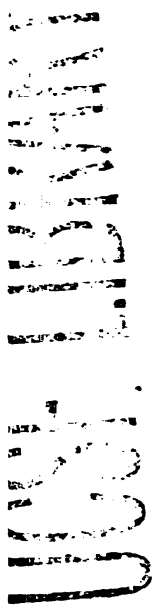
The proposed interaction between 5'SS position 2 and the 3'SS UAG could include other spliceosomal components in addition to Prp8. Excellent candidates for such an interaction are residues in the invariant ACAGAG motif of U6 snRNA. Basepairing interactions between U6 (positions 47-49) and the 5'SS consensus sequence (positions 4 through 6), (Kandels and Séraphin, 1993; Lesser and Guthrie, 1993), is thought to bring U6 G50, A51, and G52 close to 5'SS positions 1 through 3 (Table 2). Consistent with this notion, 5'SS position 2 is known to crosslink to U6 A51 (Kim and Abelson, 1996; Sontheimer and Steitz, 1993). Notably, mutations in the G50, A51, and G52 of the ACAGAG motif result in a severe block to exon ligation (Fabrizio and Abelson, 1990; Madhani et al., 1990), reminiscent of the requirements for the 5' SS GU and 3'SS UAG residues at this step.

We thus investigated whether the effect of *prp8* suppressor alleles extended to the candidate 5'SS-interacting residues in U6. We assayed the growth phenotypes of mutations in U6 G50, A51, and G52 in the presence of wild type *PRP8* or *C-122*, a representative splice site suppressor allele of *prp8* (Table 2). Recessive U6 mutant phenotypes are very severe (inviable or very sick (Madhani et al., 1990)); mutations in U6 A51, but not adjacent residues, are dominant negative (Luukkonen and Séraphin, 1998). The presence of the *prp8* suppressor allele (*C-122*), has no effect on the recessive growth phenotypes. Notably, however, the dominant negative effect of mutations in A51 is suppressed (Table 2 and Figure 6).

The ability to suppress the dominant negative effects of A51 mutants was then tested for all representative *prp8* alleles by transformation efficiency (data not shown),

growth (Figure 6A), and copper resistance conferred upon a wild type *ACT1-CUP1* (data not shown). The splice site suppressor *prp8* alleles suppress the dominant negative phenotype of A51C by these assays, while the Pyr alleles of *prp8* do not (Figure 6A and data not shown). The extent of suppression by exhibited by all of the alleles with mutations in the "B" fragment was not as strong (Figure 6A and data not shown). Thus the extent of 51C suppression appears to correlate with the location of the mutation(s) on *PRP8*.

The *prp8* suppressor alleles exerted a weak effect upon the transformation efficiency of the severely dominant negative A51U mutation, but the analysis for A51U was complicated by a high frequency of reversions (data not shown). The weak dominant effects of the A51G mutation were not detectable by growth assays. However, in primer extension assays, the *prp8* allele *C-122* was observed to allow for a small, ~%30 increase in the splicing efficiency of the wild type *ACT1-CUP1* reporter in the presence of A51G (data not shown).



The exon ligation defect is suppressed

Like the 5' splice site position 2 mutations, mutations in U6 A51 confer a strong inhibition to the exon ligation step of splicing (Fabrizio and Abelson, 1990). Thus, as for the 5' splice site mutations, the exon ligation defect of U6 A51 mutants is predicted to be suppressed by the splice site suppressor *prp8* alleles. Indeed, by primer extension analysis, we see a 2.5-fold increase in the efficiency of exon ligation for wild type *ACT1-CUP1* splicing (Figure 6B), as well as for the endogenous *RP51* transcript (data

not shown). Thus, the exon ligation defect of the dominant negative U6 mutation is partially suppressed by the *prp8* mutant.

Splice site suppressor prp8 alleles allow for assembly of A51 mutants into snRNPs

Since the *prp8* alleles did not rescue the recessive lethal phenotype of A51 mutations (Table 2), we explored the possibility that the *prp8* alleles suppress the dominant negative phenotype by impairing the ability of mutant U6 snRNA to assemble into snRNPs. This could in turn allow the assembly of more snRNPs containing wild type U6 snRNA, and thus cause an apparent suppression of the exon ligation defect. Prp8 is a component of the U5 snRNP, which joins the U4/U6 snRNP to form the U4/U5/U6 triple snRNP. It is thus possible that Prp8 forms contacts with or influences U6 in the triple snRNP. We assayed whether the *prp8* mutants affected the stability of mutant U6 snRNA, or its ability to co-immunoprecipitate with Prp8. We used strains containing pseudo-wild-type U6 (Madhani et al., 1990), which can be distinguished in size from mutant U6. The pseudo-wild-type U6 is expressed at low levels in the presence of wild-type U6, but is increased when the A51C mutant is the only other copy of U6 in the cell. As shown in Figure 6C, the *prp8* mutant *C-122* does not affect the ratio of A51C mutant U6 to pseudo-wild-type U6 in either the total pools of snRNAs, or in those immunoprecipitated by α-Prp8 antibodies. These data suggest that the *prp8* suppressor allele allows for the assembly of 51C mutant U6 into snRNPs and, probably, into spliceosomes. It is thus likely that suppression of the dominant negative phenotype of

the 51C mutation occurs on the spliceosome, at the exon ligation step. To explain why suppression of the recessive U6 A51C growth phenotype was not observed, we suggest that suppression by these *prp8* alleles is simply not strong enough to confer viability to the severe U6 mutant.

Mutations adjacent to U6 A51 are not suppressed

Within U6, suppression by the *prp8* alleles appears specific to U6 A51. Combining the *prp8* alleles with mutations in U6 nucleotides adjacent to A51 did not result in any growth alterations (Table 2), despite the fact that mutations in G52 and G50, like A51, have been shown to strongly impair exon ligation (Fabrizio and Abelson, 1990). For a more quantitative comparison, we conducted the primer extension analysis of *ACT1-CUP1* splicing for the viable U6 G52 and G50 mutations. In contrast to what is observed with A51C, neither the efficiency of exon ligation, nor the overall splicing efficiency, was improved (data not shown). Thus, while a number of U6 mutations impede exon ligation, suppression by *prp8* appears specific for mutations in A51.

DISCUSSION

A new tertiary interaction between the 5'SS, 3'SS and U6

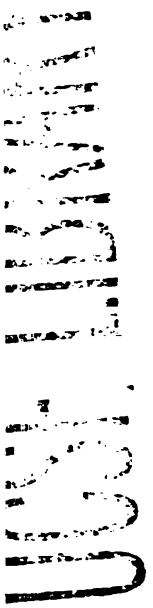
Our initial goal was to provide functional correlates for Prp8's crosslinking interactions to the 5' and 3' splice site consensus sequences, and to determine whether functions at the two splice sites would map to distinct or to the same functional domains

of the protein. We found *prp8* alleles that suppress mutations in both the 5'SS GU and in the 3'SS UAG; indeed, all *prp8* alleles that suppress mutations in the 3'SS UAG also suppress mutations in 5'SS position 2. Furthermore, we found that mutations in A51 of the U6 ACAGAG motif are also suppressed by the *prp8* alleles. However, other mutations in intron consensus sequences, and in neighboring residues of U6, are not suppressed. To explain the distinct, but unanticipated pattern of suppression, we suggest that Prp8 influences a specific, previously unrecognized, tertiary interaction between the suppressed residues (Figure 7A,C)

This proposed interaction is consistent with several previous observations. First, mutations in the 5' and 3' terminal intron residues can reciprocally suppress each other, with a specificity that suggests that they form a non-Watson-Crick basepairing interaction (Chanfreau et al., 1994; Deirdre et al., 1995; Parker and Siliciano, 1993). Second, 5'SS position 2 crosslinks to U6 A51 (Kim and Abelson, 1996; Sontheimer and Steitz, 1993). In addition to Watson-Crick pairing of U6 with positions 4 through 6 of the 5'SS (Kandels and Séraphin, 1993; Lesser and Guthrie, 1993), genetic interactions between U6 and 5'SS positions 1 and 3 support the hypothesis that the entire 5'SS consensus sequence is juxtaposed to the U6 ACAGAG sequence (Luukkonen and Séraphin, 1998). Third, the U6 G52U mutation suppresses mutations in the terminal intron residue, suggesting that U6 and the 3'SS interact at least indirectly (Lesser and Guthrie, 1993).

Importantly, while all of the above previous observations are consistent with an extended interaction between the 5' and 3' ends of the intron (Figure 7A), they have been equally consistent with a configuration (in Figure 7B) in which the sole interaction

between the 5' and 3' ends of the intron occurs between the first and last nucleotides. Evidence for an interaction between the second and penultimate nucleotides of the intron, predicted by Figure 7A, has not been found by directed mutagenesis (Ruis et al., 1994). The tertiary interaction suggested by our data between 5'SS position 2 and the 3'SS UAG now favors the extended interaction between the 5' and 3' ends of the intron, (Figure 7A, C). An interaction between the second and penultimate residues of the intron is thus likely to occur, and could have eluded previous studies through constraints imposed by additional factors that were not mutagenized in that study, namely U6 and Prp8. Supportive of this idea, an indirect interaction was suggested to occur between 5'SS position 3, and 3'SS position -3, based on the observation that mutations in 5'SS position 3 nonspecifically affect the competition of closely spaced 3'SS UAG sequences (Deirdre et al., 1995). Prp8 and U6 may also affect a tertiary interaction between these positions (see below).



It is striking that participants in the postulated interaction are conserved between the conventional U2-dependent and the divergent U12-dependent spliceosomes. Although introns spliced by the two different pathways vary in whether they contain G-G or A-C in the first and last positions, all contain U in the second and A in the penultimate positions (Dietrich et al., 1997; Sharp and Burge, 1997; Shukla and Padgett, 1999). Moreover, the residue in U6atac snRNA corresponding to position 51 of yeast U6 is also an A (Tarn and Steitz, 1996). Additionally, the three adjacent residues: 5'SS position 3, 3'SS position -3 of the intron, and the position corresponding to U6 G50, are also conserved (Dietrich et al., 1997; Sharp and Burge, 1997; Tarn and Steitz, 1996). Notably, Prp8 has also been shown to be a component of the two spliceosomes (Luo et al., 1999). This supports

the idea that the proposed interaction of the three RNAs and Prp8 is fundamental to the mechanism of pre-mRNA splicing.

Interactions that bridge the two catalytic steps of splicing

The longstanding observation that residues at the 5' end of the intron are required for the second transesterification step has always seemed paradoxical, since the 5' end of the intron is no longer a substrate for splicing chemistry after the first step is completed. Indeed, in the splicing reaction catalyzed by group I introns, the 5' end of the intron leaves the active site after the first cleavage step, in order to be replaced by the 3' end of the intron for the second transesterification step (Cech, 1990). In the spliceosome, the 5' end of the intron could function at the active site for both steps. If the 5' exon does not stray far from the cleaved 5' end of the intron, these residues could serve, through the proposed interaction (Figure 7C), to position the 3' splice site for catalysis of exon ligation.

Little is known about how the spliceosome couples 5' splice site cleavage to exon ligation. The 5' exon is thought to be held or carried from the first to the second catalytic step by an invariant loop in U5 snRNA (reviewed in Newman 1997). Because Prp8 makes extensive crosslinks to the loop, and to the 5' and 3' exons, it has been hypothesized that Prp8 acts in conjunction with the U5 loop to position the exons for ligation (Dix et al., 1998; O'Keefe and Newman, 1998; Teigelkamp et al., 1995). Alignment of the 5' and 3' ends of the intron could help in the alignment of the exons and in the

positioning of the 3' splice site for catalysis (Figure 7C). This could be accomplished by the proposed network of RNA interactions and facilitated by Prp8.

The progression of splicing from 5' splice site cleavage step to exon ligation is known to involve a conformational change in the spliceosome (Chiara et al., 1996; Chua and Reed, 1999; Schwer and Guthrie, 1992; Umen and Guthrie, 1995) and the construction of a chemically distinct active site (Moore and Sharp, 1993; Sontheimer et al., 1997). However, a number of interactions that are critical for exon ligation are established during spliceosome assembly, before 5' splice site cleavage. These include the interaction of the 5'SS consensus sequence with the U6 ACAGAG motif (Kim and Abelson, 1996), and the 5' exon with the U5 invariant loop (Newman and Norman, 1991; Newman and Norman, 1992; O'Keefe et al., 1996; Sontheimer and Steitz, 1993; Wassarman and Steitz, 1992). Prp8 is also observed to crosslink to the 5'exon (Chiara et al., 1996; Teigelkamp et al., 1995; Wyatt et al., 1992) and to the 5'SS GU before 5' cleavage (Reyes et al., 1996). In a concurrent study, Siatecka *et al.*, report evidence that Prp8 functions in recognition of 5'SS position 2 during spliceosome assembly, before 5'SS cleavage, as well as during exon ligation (Siatecka et al., 1999). The maintenance of interactions between the two catalytic steps of splicing suggests ways in which one active site could be altered to undergo both transesterification steps. Through interactions with the 5'SS, U6 and 3'SS, in addition to the 5'exon, U5, and 3' exon, Prp8 could perform a critical function in bridging the two steps of splicing.

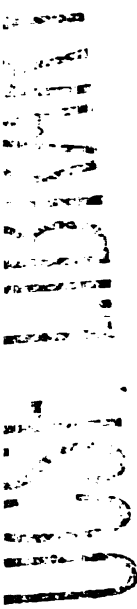
In another recent study (Kuhn and Brow, 1999), a mutation in *PRP8, (prp8-201),* was found to suppress a U4 mutation, U4cs1, which compromises the interaction of the U6 ACAGAG motif with the 5'SS consensus sequence. Interestingly, suppression is

conferred by a mutation in residue 1861 of Prp8, which is strikingly close to the location of mutations in the splice site suppressor alleles *D-135, D-136*, and *D-143* (Table 1). In fact, *prp8-201* suppresses 5'SS U2A and 3'SS UAG mutations (C. C. and C.G., unpublished results). However, our splice site suppressor alleles fail to suppress U4cs1 (A.K and D.B., personal communication); thus, the relationship between splice site suppression and U4 cs1 suppression by *prp8* remains to be understood. Notably, U4cs1 suppression appears to occur prior to 5' cleavage (Kuhn and Brow, 1999). The fact that one mutation, in *prp8-201,* can suppress defects at two different steps suggests, again, that Prp8 forms interactions that are maintained throughout both chemical steps of splicing.

A role for Prp8 at the spliceosomal catalytic core

The model based on our genetic results provides an attractive explanation for the extensive crosslinking interactions that have previously been observed between Prp8 and residues near both the 5' splice site and 3' splice site. These data suggest that Prp8 actually binds the proposed tertiary RNA interaction. However, it is not known whether the amino acids mutated in the suppressor alleles, which map to four distinct regions of the protein, play a direct role in RNA recognition. While most of these residues are conserved in the known homologues of Prp8, all of the suppressor mutations are viable, and confer only a mild, if any, detectable defect to wild type splicing. The suppressor alleles could serve to relax some of the constraints imposed upon the identity of splice site residues, thereby allowing mutant as well as wild type sequences to be utilized. In

well-studied systems of protein-substrate recognition, mutations that permit a wider range of interactions have been observed to increase the flexibility of the binding interface or pocket (Bone et al., 1989; Morton and Matthews, 1995). Such effects are not limited to mutations at the site of binding, and, in fact, may occur at considerable distances (Gutierrez et al., 1998; Mace et al., 1995). Indeed, it remains formally possible that the suppression observed is mediated by allosteric effects upon another protein that binds the 5' and 3' ends of the intron. However, despite the numerous proteins that have been found to function during the exon ligation step (Umen and Guthrie, 1995), Prp8 is the only protein that has yet been found to affect the requirement for UAG at the 3' splice site.



While spliceosome catalysis is thought to be performed by RNA, many of the same types of observations that suggest a critical role for RNA at the catalytic core have yielded similar results for Prp8, namely, crosslinking to RNA near sites of chemistry, genetic suppression of splice site mutations, and phylogenetic conservation. What is, then, the specific role of Prp8? Prp8 could conceivably make direct structural or chemical contributions to the spliceosomal active site. Alternatively, Prp8 could serve to add constraints or stability to a structure that is intrinsically comprised of RNA. This latter role for protein in RNP catalysis has been demonstrated for group I introns (Weeks and Cech, 1996; Weeks and Cech, 1995; Weeks and Cech, 1995) and Rnase P (Guerrier-Takada et al., 1983). Recent crystal structures of ribosomal L11 protein complexed with a 58-nt domain of 23S rRNA beautifully illustrate that a protein component of an RNP can stabilize an unusual RNA fold through direct contact with a critical RNA tertiary interaction (Conn et al., 1999; Wimberly et al., 1999). Ultimately, an ultra-structural

analysis of Prp8's interaction with RNA in the spliceosome is required to understand the contribution of this protein to catalysis.

MATERIALS AND METHODS

Yeast strains and plasmids

Screens and characterization of *prp8* alleles were conducted with the strain yJU75 (Umen and Guthrie, 1996): MATa, *ade2 cup1Δ::ura3 his3 leu2 lys2 prp8Δ::LYS2 trp1*; pJU169 *(PRP8 URA3 CEN ARS). ACT1-CUP1* reporters are described in (Lesser and Guthrie, 1993) and (Umen and Guthrie, 1996). For all reporters, the *LEU2*-marked (pGAC24) version was used. The 5'SS U2A, U2G, and wild type *ACT1-CUP1* reporters are described in (Ruis et al., 1994), except that the pGAC14 vector backbone was swapped for pGAC24 (Lesser and Guthrie, 1993), to generate pCC72 (U2A), pCC44 (U2G), and pCC71 (wt).

For testing genetic interactions between *prp8-122* and mutations in U6 snRNA (Figures 6 and 9), a *prp8Δ::LYS2* disruption was generated, according to (Umen and Guthrie, 1996), in a diploid heterozygous for the disruption *snR6Δ::LEU2* (Brow and Guthrie, 1988). The diploid was then transformed with plasmids (see below) and sporulated. Haploid progeny containing both disruptions: *prp8Δ::LYS2* covered by wild type *PRP8* or *prp8-C122* on a pRS423 plasmid *(2 micron, HIS3)*, and *snR6Δ::LEU2* covered by wild type U6 in pSE360, were identified by nutritional markers, and the presence of both *prp8* and *snR6* disruptions was confirmed by whole cell PCR. The

resulting U6-shuffle strains were transformed with plasmids containing wild type (pSX6T) or mutant U6 (Madhani et al., 1990), and assayed for growth before or after counter-selection on 5-FOA of the *URA3*-marked plasmid containing wild type U6 (Table 2). For the experiments in Figure 6C, a similar strategy was used to generate a double knockout haploid strain containing plasmid-borne pseudo-wild type (pwt) U6, and wild type or mutant *prp8* (in pRS423). The resulting strain was then transformed with plasmids (pSE360) containing wild type or A51C mutant U6.

Screens for suppressors of mutations in the 5'SS GU

PRP8 was PCR-mutagenized in four parts, using primers and Mn^{2+} conditions described in (Umen and Guthrie, 1996), according to mutagenic PCR conditions described in (Leung et al., 1989). *prp8* mutants consisting of each of four mutagenized fragments were created by *in vivo* gap repair (Muhlrad et al., 1992). The four corresponding gapped plasmids were generated from pJU225 *(PRP8 TRP1 2μ)*, as described in (Umen and Guthrie, 1996).

yJU75 strains containing either 5'SS U2A or U2G mutant *ACT1-CUP1* reporters were transformed, and, for each reporter, approximately 4,000 were replica-plated onto media containing 0.025, 0.05, or 0.1 mM copper sulfate (Lesser and Guthrie, 1993). In order to look for recessive as well as dominant suppression, transformants that survived passage on 5-FOA were re-screened for increased growth on copper. For each candidate suppressor, the *TRP*-marked plasmid was recovered, transformed into a fresh yJU75 strain, and retested for the ability to confer increased growth on copper with the 5'SS U2A and U2G *ACT1-CUP1* reporters.

Copper growth

Duplicate cultures were grown to log phase in media lacking leucine to ensure maintenance of the *ACT1-CUP1* reporter plasmid. Cultures were then diluted to an OD_{600} of 0.5, and a frogger was used to stamp equivalent amounts of cells from each strain onto plates containing different concentrations of copper (Lesser and Guthrie, 1993). For assays of the effect of dominant negative mutations in U6 A51 (Figure 6A), cells were grown on media and copper plates lacking histidine, to ensure maintenance of the pSE362 plasmid containing the dominant negative U6 mutant.

Mapping of suppressor mutations

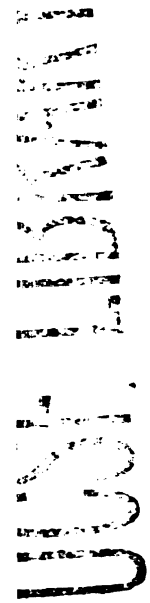
Since each mutagenized *prp8* fragment was approximately 2kb, and contained multiple mutations, the location of mutations sufficient to confer suppression was mapped to a smaller region of *PRP8*. Fragments of the mutant *prp8* were PCR-amplified in non-mutagenic conditions with Pfu polymerase. The PCR fragments were then co-transformed with appropriately gapped plasmid. The fragments of mutant *prp8* coding sequence that could reconstruct the suppressor phenotype for each mutants is listed:

Allele	Fragment PCR-amplified from mutant:	Enzyme used to gap pJU225:
C-133, C-134	4262-5161	*BstE II* (w. EtBr)*
D-135, D-136, D-143	5340-6147	*Msc I*
D-144	6017-7215	*Stu I*
B-131, B-132, B-141, B-142	1635-3322 or 2431-3735	*Afl II* (w. EtBr)* *Spe I*

Oligos used to amplify each fragment ranged in size from 20 to 22 nucleotides. *For enzymes with two sites in pJU225, the plasmid was partially digested in the presence of 100ug/ml ethidium bromide to enrich for singly cut plasmids (Umen and Guthrie, 1996).

The region listed for each allele was then sequenced at the Biomolecular Resource Center DNA Sequencing facility at UCSF, using the "Big Dye" Terminator Cycle Sequencing Kit (Applied Biosystems). The program ALIGN (Myers and Miller, 1988) was used at the GENESTREAM network server (http://www2.igh.cnrs.fr) to identify changes in the *PRP8* sequence.

The three mutations in the sequenced region of *D-144* were isolated from each other by standard cloning techniques, using the intervening *Stu I* and *Sph I* sites. Each isolate was confirmed by sequencing, and found to confer modest suppression of 5'SS U2A. Likewise, *D-134* and *D-143* mutations were separated from additional mutations using the *Msc I* site. These additional mutations (not shown) were not required for the suppression phenotype.



Primer extension analysis

RNA preparation and primer extension assays were performed as previously described (Frank and Guthrie, 1992; Lesser and Guthrie, 1993). U14 snoRNA was primer extended for an internal control for the amount of total RNA in each lane (Noble and Guthrie, 1996). Products were quantified by phosphorimager analysis of duplicate or triplicate samples. The total splicing efficiency and the efficiency of the exon ligation step were estimated as previously derived (Fouser and Friesen, 1986; Pikielny and Rosbash,

1985), by calculating the ratio of levels of mature/precursor species for the total splicing efficiency, and mature/lariat for the efficiency of the exon ligation step.

Bead-beat extracts

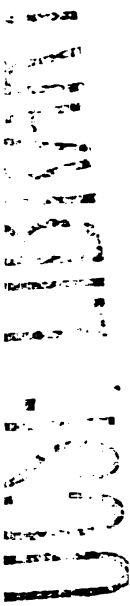
For figure 6C, strains described above were grown to late log phase in media lacking uridine, to ensure maintenance of the dominant negative U6 plasmid. 100 mls of culture were pelleted and resuspended in 150 μl of Lysis Buffer (45mM Hepes pH 7.9, 400 mM NaCl, 0.5% NP-40 (vol./vol.), 1 mM EDTA, 20% glycerol (vol./vol.), DEPC-treated dH20, and protease inhibitors (1 mM PMSF, 1ug/ml leupeptin, 1 mM benzamidine)). Cells were lysed by vortexing with 600 μl acid-washed glass beads (size 0.5 mm), in 1 minute bursts followed by 1 minute rest on ice, repeated 6 times. Insoluble material in the lysate was then removed by microfuging for 10 minutes at maximum speed. The supernatant lysate was the diluted with an equal volume of Dilution Buffer (45 mM Hepes pH 7.9, 1mM EDTA, 20% glycerol, with a fresh addition of protease inhibitors). All steps were conducted at 4°C.

Polyclonal α-PRP8 antibodies

The lacZ-PRP8 fusion plasmid, pFP8.4, was provided by Jean Beggs, and expression of the fusion protein was heat induced in the E. coli strain pop2136 as previously described (Lossky et al., 1987). The insoluble pellet, which contained the majority of the fusion protein, was prepared from lysed cells, and purified by according

to (Harlow, 1988). The ~170 kDa fusion protein was further purified by subjecting the sample to SDS polyacrylimide gel electrophoresis, and removing the band after staining with 0.05% Coomassie Blue in distilled water. Gel slices, containing a total of ~0.5 mgs of fusion protein, were solubilized in Freunds adjuvant, and used to immunize rabbits. The rabbits were boosted in three-week intervals with 0.25 mgs of gel-purified fusion protein. Pre-immune sera were collected prior to immunization, and antisera were obtained ~10 days after the fourth boost. All rabbit work was done by Berkeley Antibody Company (BabCO) in Richmond, CA.

The ability of the antibodies to immunoprecipitate Prp8 was confirmed using extracts from strains that contain hemagglutinin (HA) epitope-tagged *PRP8* (Umen and Guthrie, 1995), and subjecting the pellets and supernatants to Western Blot analysis with a monoclonal antibody specific for the HA epitope.



Immunoprecipitations

For immunoprecipitation of snRNAs, IgG was pre-bound to 30 μl Protein-A Sepharose (Sigma) in NTN buffer (150mM NaCl, 50 mM Tris-HCl pH7.4, 0.1% NP-40 (vol./vol.)), and washed five times with 600 μl of the same buffer at 4°C. For each experiment, 200 μl of bead-beat extract was added to 15μlof pre-bound serum, and incubated at 4°C, with protease inhibitors (above), for 2.5 hours. The antibody complexes were centrifuged, and washed four times with 600 μl NTN buffer. RNA was then prepared from the pellets, (and 10 μl of total extract), and subjected to Northern analysis, using a ^{32}P kinased oligo probe against U6 snRNA (Noble and Guthrie, 1996)

incubated at 4°C, with protease inhibitors (above), for 2.5 hours. The antibody complexes were centrifuged, and washed four times with 600 µl NTN buffer. RNA was then prepared from the pellets, (and 10 µl of total extract), and subjected to Northern analysis, using a ^{32}P kinased oligo probe against U6 snRNA (Noble and Guthrie, 1996)

ACKNOWLEDGMENTS

We thank Andreas Kuhn and Dave Brow for communicating unpublished results; Magda Konarska, Mirka Siatecka, and José Reyes for communicating results and for extensive collaborative discussions; Jim Umen, Jon Staley, Stephen Rader and David McPheeters for helpful discussions; Stephen Rader, Amy Kistler, Kent Duncan, Maki Inada, Jon Staley, James Wilhelm, Alan Frankel, and John Abelson for critical comments on this manuscript; and Lucita Esperas and Carol Pudlow for excellent technical assistance.

This chapter was published in *Genes and Development*, vol. 13: 1970-82.

TABLE 1

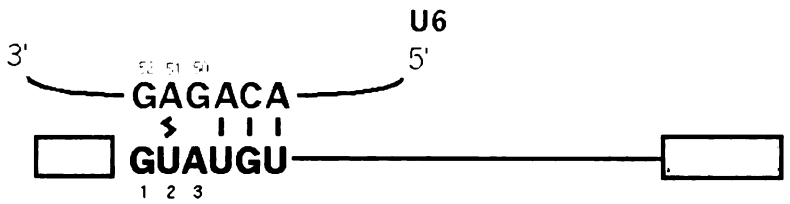
Identity of 5'SS position 2 suppressor mutations in prp8

PRP8 allele	Mutation(s)
B-131	Y923C, E941G
B-132	I857T, E888G, Y923C, E935G, E941G, Q976R, D1075G
C-133	I1444V, T1565A, V1621A
C-134	Q1455P, D1485G, H1592R
D-135	**E1817G**
D-136	**N1869D**
B-141	I857T, Y923C, E941G
B-142	I857T, S894G Y923C, E941G, S1018P
D-143	**K1864E**
D-144	F2176S*, Q2313R*, T2364A*

Most of the suppressors contained additional mutations, not listed, which did not affect the suppression phenotype (see methods). Single mutations that confer the suppression phenotype are indicated in bold. **D-144* contains 3 mutations, each of which confer a weak suppression phenotype on their own, but confer stronger suppression when combined. For other alleles, the mutations listed have not been separated from each other to identify those responsible for suppression.

TABLE 2

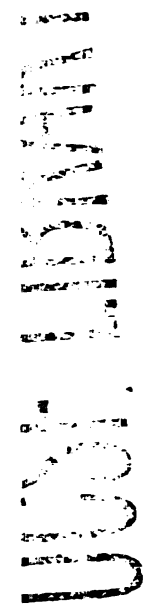
Genetic interactions between a *prp8* splice site suppressor allele and U6 snRNA



	(+ wt U6)		(- wt U6)	
	wt PRP8	*mut prp8*	*wt PRP8*	*mut prp8*
wt U6	+++	+++	+++	+++
G50A	+++	+++	--	--
C	+++	+++	+	+
U	+++	+++	+	+
A51C	+	++	--	--
G	+++	++	--	--
U	--	+	--	--
G52A	+++	+++	--	--
C	+++	+++	--	--
U	+++	+++	+	+

Mutations in U6 snRNA residues of the ACAGAG motif were tested for genetic interactions with the prp8 suppressor allele, *C-122*. Growth (at 30°C) was assayed in either in the presence or absence of an extra copy of wild type U6. The effects conferred by *C-122* upon mutant U6 are shaded. (--) no growth. (+) or (++) intermediate growth. (+++) wild type growth rate. *While effects on growth were not detected, a 30% increase in splicing efficiency was measured for a wild type *ACT1-CUP1* reporter by primer extension assays.

Figure 1: Prp8 forms extensive crosslinks to RNA residues in the intron consensus sequences, and in the 5' and 3' exons. Intron consensus sequences that define the 5' splice site (5'SS), branchpoint (BP), and 3' splice site (3'SS) are indicated. For most yeast introns, a stretch of pyrimidines (Pyr), functions with the 3'SS UAG in definition of the 3' splice site. Triangles denote the location of Prp8 crosslinks observed in either mammalian (solid), or yeast (outline) systems. For all studies except (*), the location of the crosslinked site is determined through the use of a site specific photo-reactive substitution, or by RNase fingerprinting. (*)s represent approximated sites of crosslinking within an RNase T1 fragment.



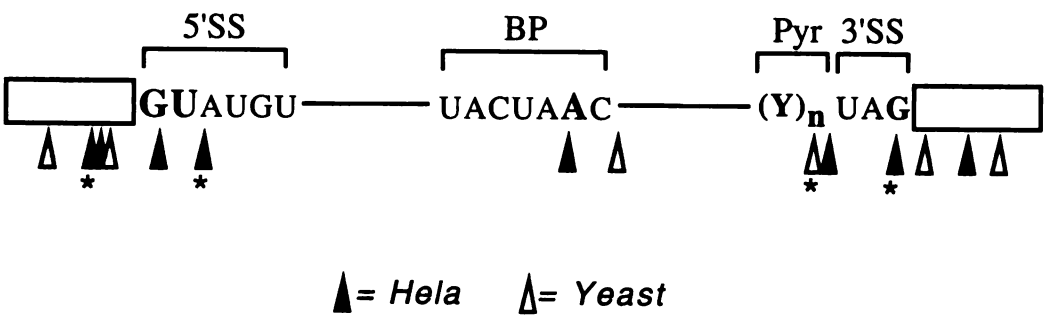
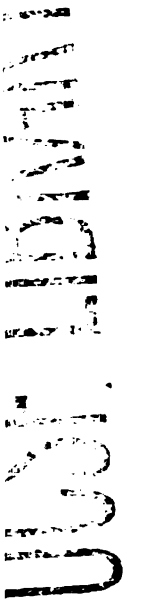

5'SS
BP
Pyr
3'SS
GUAUGU
UACUAAC
(Y)n UAG
▲= Hela
△= Yeast

Figure 2: Identification of *prp8* alleles that suppress mutations in position 2 of the 5' splice site.

A. Strategy for isolation of new *prp8* alleles. The coding region of *PRP8* was divided into four equal fragments (A,B,C,D). Each fragment was amplified by mutagenic PCR and co-transformed with an appropriately gapped plasmid (as shown for B), to generate mutant alleles of *prp8* by *in vivo* gap repair. The resulting strain, which is deleted for its chromosomal copy of *PRP8* and *CUP1*, is diagrammed. Mutant *prp8* transformants were screened for the ability to confer increased copper growth upon *ACT1-CUP1* reporters containing U2A or U2G mutations in the 5' splice site, both before and after loss of the wild-type *PRP8* allele by 5-FOA selection.



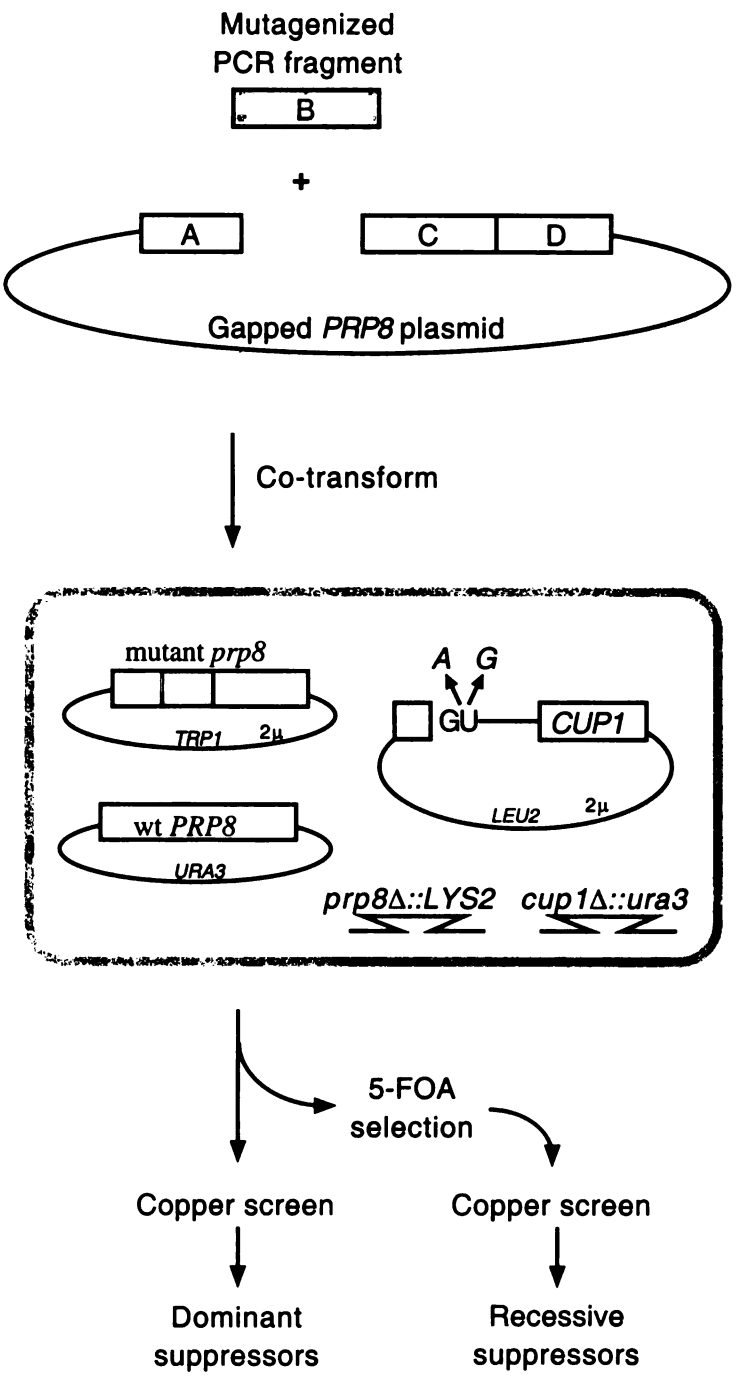
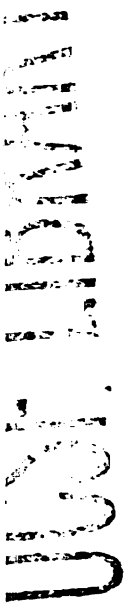

Mutagenized
PCR fragment
B
+
A
C
D
Gapped PRP8 plasmid
Co-transform
mutant prp8
TRP1
2μ
wt PRP8
URA3
A G
GU
CUP1
LEU2
2μ
prp8Δ::LYS2
cup1Δ::ura3
5-FOA
selection
Copper screen
Copper screen
Dominant
suppressors
Recessive
suppressors

Figure 2: Identification of *prp8* alleles that suppress mutations in position 2 of the 5' splice site.

B. Suppression of the copper growth phenotype of mutations in 5'SS position 2 by the newly identified *prp8* alleles. Each column shows growth on the concentration of copper, 0.075 mM and 0.05 mM, that is limiting for 5'SS U2A and U2G reporters, respectively. Equal numbers of log phase cells from each strain were spotted onto the same copper containing plate.



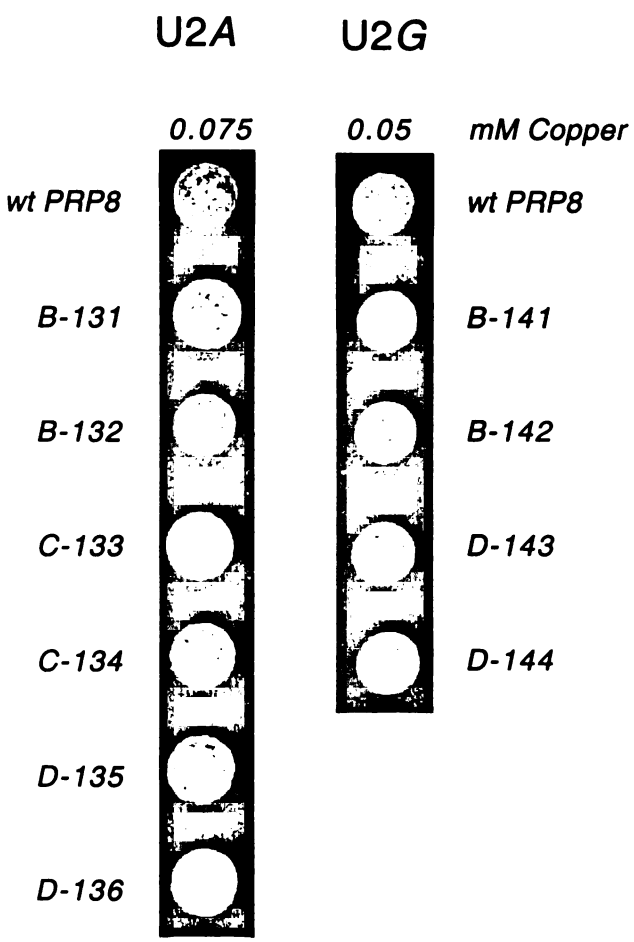
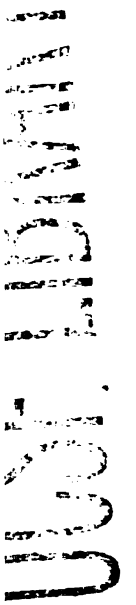

U2A
U2G
0.075
0.05
mM Copper
wt PRP8
B-131
B-132
C-133
C-134
D-135
D-136
wt PRP8
B-141
B-142
D-143
D-144

Figure 3: 5'SS position 2 suppressor mutations cluster in four regions of *PRP8*. The location of mutations that confer 5'SS position 2 suppression are indicated in black, while green and blue indicate the location of previously characterized 3'SS UAG suppressors and Pyr alleles, respectively (Umen and Guthrie, 1996). The corresponding location of the 5'SS GU crosslink in mammalian Prp8 is indicated by the lightning bolt. Solid lines indicate the location of single mutations that confer suppression. Dotted lines indicate mutations which confer suppression together, but which have not been assayed on their own.



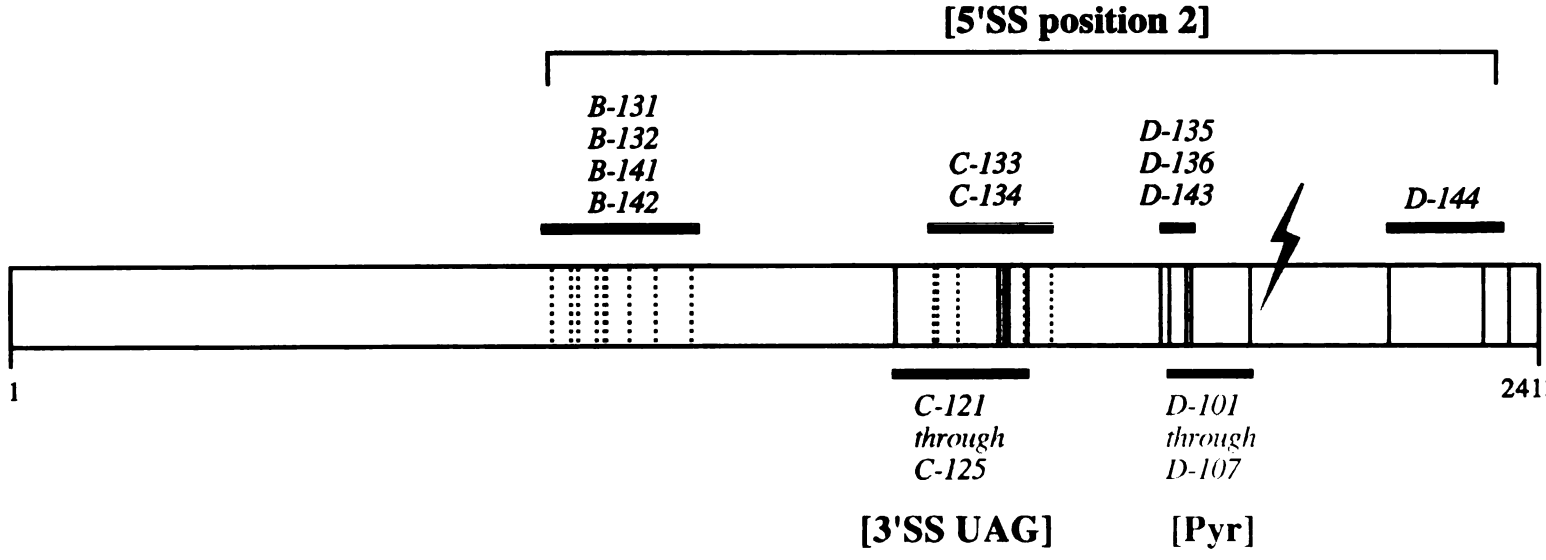
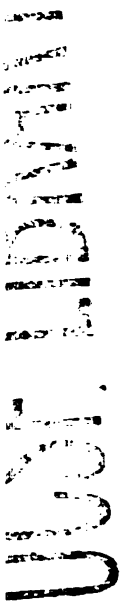

[5'SS position 2]
B-131
B-132
B-141
B-142
C-133
C-134
D-135
D-136
D-143
D-144
1
2413
C-121
through
C-125
D-101
through
D-107
[3'SS UAG]
[Pyr]

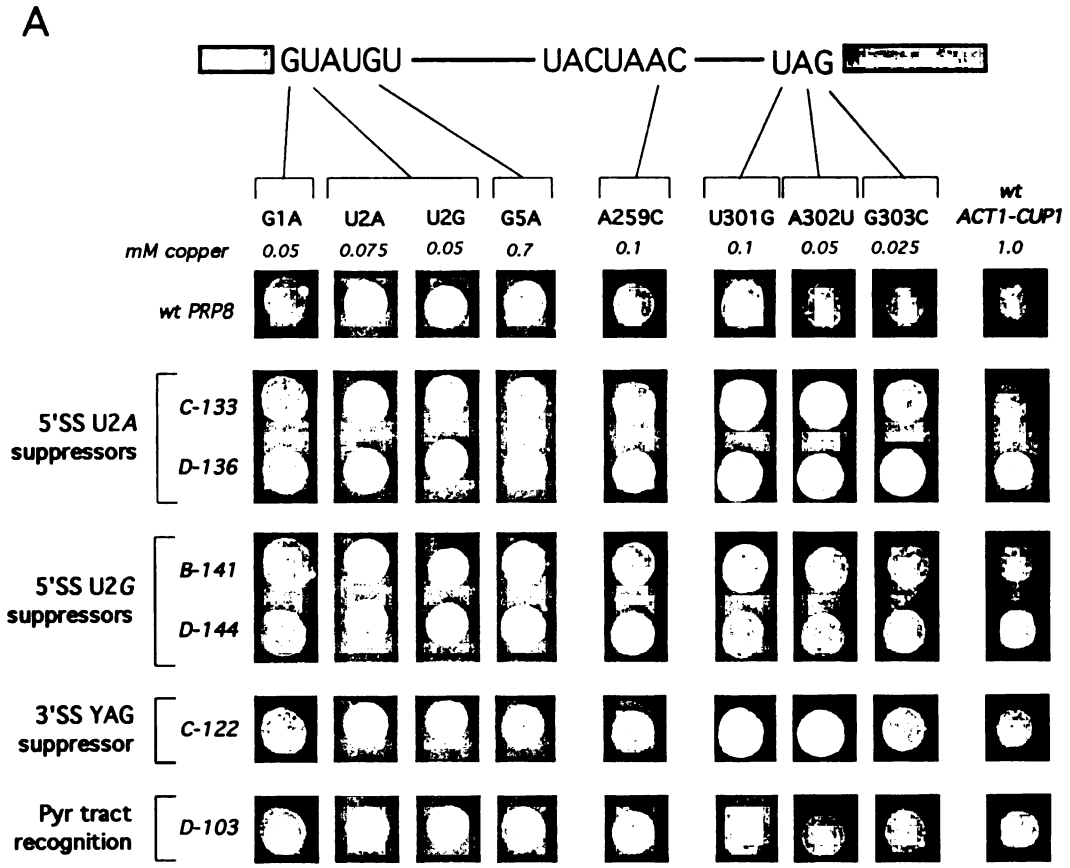
Figure 4: Allele specificity analysis reveals overlapping phenotypes at the 5' and 3' splice sites.

A. *prp8* alleles identified in different screens were tested for effects upon reporters containing various mutations at the 5' splice site, branchpoint, or 3' splice site. Each column shows growth on the concentration of copper that is limiting for each reporter in a wild type *PRP8* strain. Equal numbers of log phase cells for each mutant strain were spotted onto the same copper containing plate. The *prp8* alleles do not confer strong effects on growth when the *ACT1-CUP1* reporter is wild type, (*C-133*, the exception, is temperature sensitive). Not visible, some but not all of the alleles from each of the screens were observed to confer a mild enhancement of copper resistance to the wild type *ACT1-CUP1* reporter. Because some Pyr alleles display this puzzling effect, it does not appear to correlate with the 5'SS and 3'SS suppression phenotype.



B. Phenotypes for *prp8* alleles identified in three independent screens are summarized. "+" indicates intron mutations that are suppressed, while "-" indicates intron mutations that are not suppressed. Alleles identified through independent screens for 5'SS position 2 and 3'SS UAG suppression display the same "splice site suppression" phenotype.

A

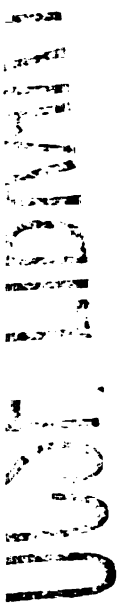


B

	5'SS				BP	3'SS			wt
	G1A	U2A	U2G	G5A	A259C	U301G	A302U	G303C	ACT1-CUP1
Splice site suppressors 5'SS:	-	+	+	-	-	+	+	+	(-)
Splice site suppressors 3'SS:	-	+	+	-	-	+	+	+	(-)
Pyr:	-	-	-	-	-	-	-	-	(-)

Figure 5: The exon ligation defect of 5'SS U2A and U2G splicing is suppressed.

A. Primer extension analysis of *ACT1-CUP1* RNA splicing. Products generated from precursor, mature, lariat-intermediate, and U14 control species are denoted on the left of the gel. Other bands are prominent primer extension stops derived from the longer precursor species. To the right, splicing of a wild type *ACT1-CUP1* reporter is shown, to highlight the splicing defect conferred by the U2A mutation and the modest extent of suppression by *prp8* alleles. Note that this experiment used less total RNA, as indicated by the reduced U14 levels.



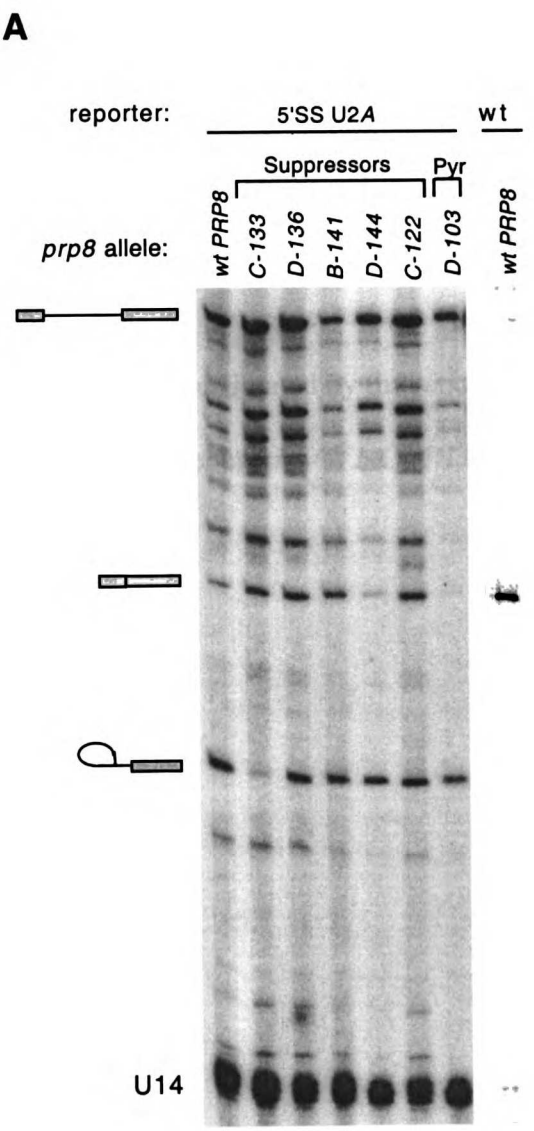
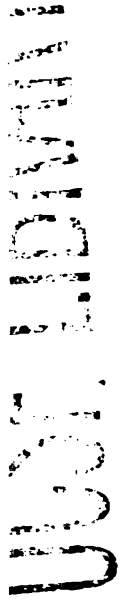


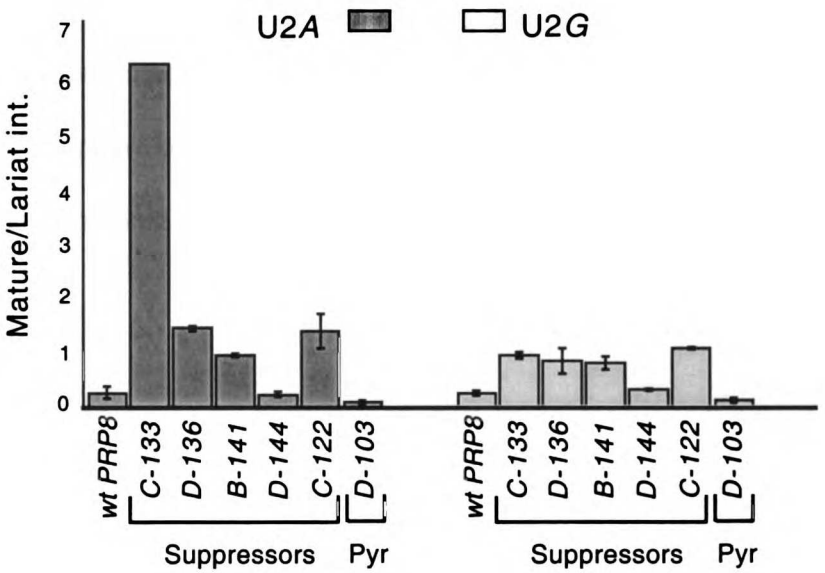
Figure 5: The exon ligation defect of 5'SS U2A and U2G splicing is suppressed.

B. The efficiency of the second step is estimated by calculating the ratio of levels of mature spliced *ACT1-CUP1* RNA to the lariat intermediate species (Fouser and Friesen, 1986; Pikielny and Rosbash, 1985). The values were determined through phosphorimage analysis of triplicate samples, with the exception of *C-133* (for U2A). The near background levels of lariat intermediate in this strain resulted in large deviations for the (Mat/Lar int.) measurement. Shown is the lowest estimate for this ratio.

C. The total splicing efficiency is estimated by the ratio of levels of mature spliced *ACT1-CUP1* RNA to unspliced precursor *ACT1-CUP1* RNA.



B



C

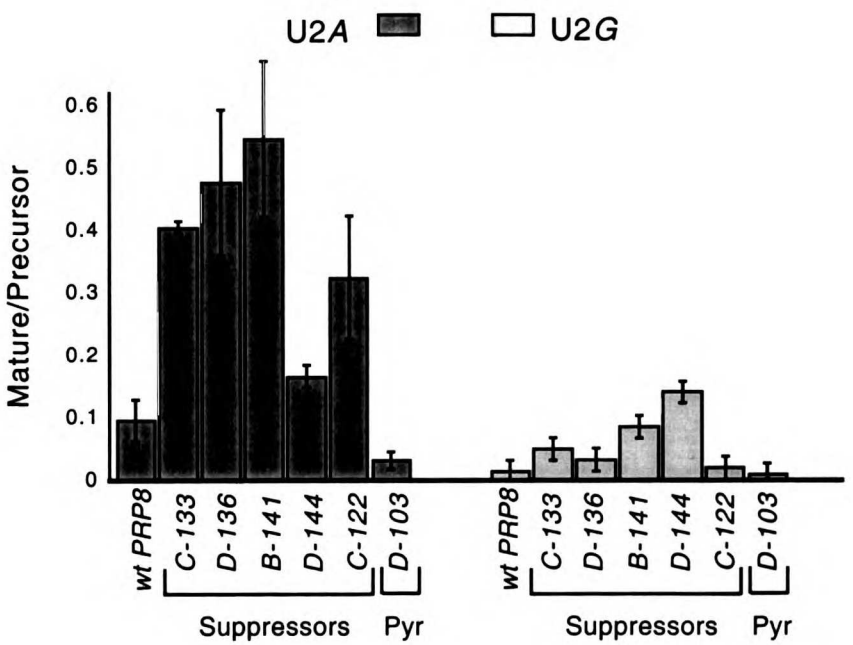
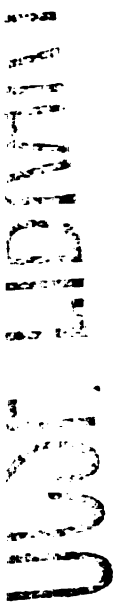


Figure 6: Splice site suppressor alleles of *prp8* suppress the dominant negativity of mutations in U6 A51.

A. Growth (at 30°C) conferred by each *prp8* allele in strains containing wt or 51C mutant U6 on a plasmid, in addition to a wild type chromosomal copy of U6 snRNA. Equal numbers of log phase cells were spotted onto medium lacking histidine to require retention of the *HIS3*-marked plasmid containing dominant negative U6.



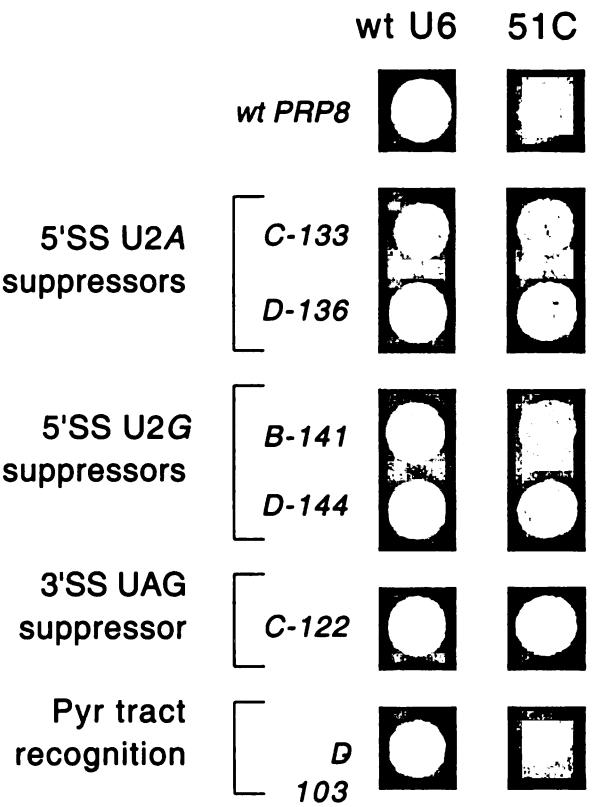
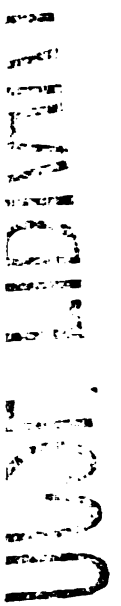

wt U6
51C
wt PRP8
5'SS U2A suppressors
C-133
D-136
5'SS U2G suppressors
B-141
D-144
3'SS UAG suppressor
C-122
Pyr tract recognition
D
103

Figure 6: Splice site suppressor alleles of *prp8* suppress the dominant negativity of mutations in U6 A51.

B: The exon ligation defect conferred by dominant negative mutations in U6 A51 is partially suppressed by *prp8* alleles. Primer extension analysis of wild type *ACT1-CUP1* reporter RNA was conducted in triplicate for wild type *PRP8* and *C-122* strains. The efficiency of the exon ligation step of splicing was estimated as described in Figure 5.



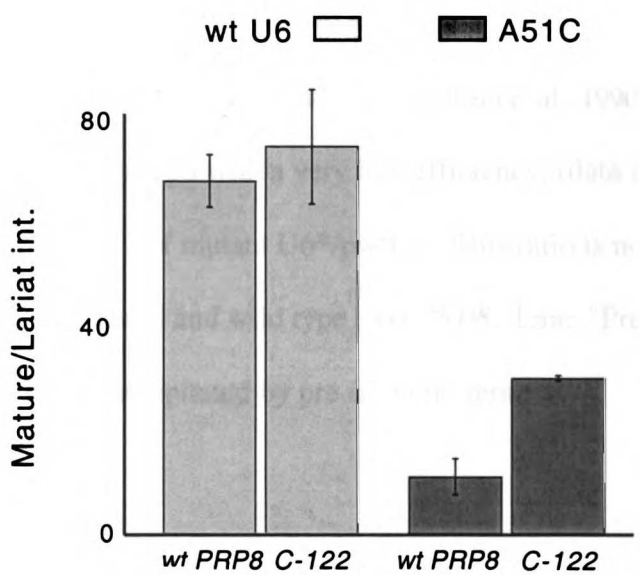


Figure 6: Splice site suppressor alleles of *prp8* suppress the dominant negativity of mutations in U6 A51.

C. *prp8* suppressor allele *8-C-122* allows for expression and association of U6 A51C. Strains containing mutant or wild type *prp8*, and a shorter, pseudo-wild-type version of U6 snRNA (pwtU6), were transformed with an additional plasmid containing full length U6*: either wild type or A51C. Total RNA and RNAs that co-IP with polyclonal α–Prp8 were resolved on a denaturing gel and subjected to Northern analysis with an oligo probe that hybidizes to both pwtU6 and U6*(Madhani et al., 1990). The polyclonal antibodies immunoprecipitate Prp8 with a very low efficiency, (data not shown), but the critical comparison is the ratio of mutant U6*/pwtU6. This ratio is not significantly different between *8-C-122* (*mut*) and wild type (wt) *PRP8*. Lane "Pre" shows the background levels of RNA precipitated by pre-immune sera.

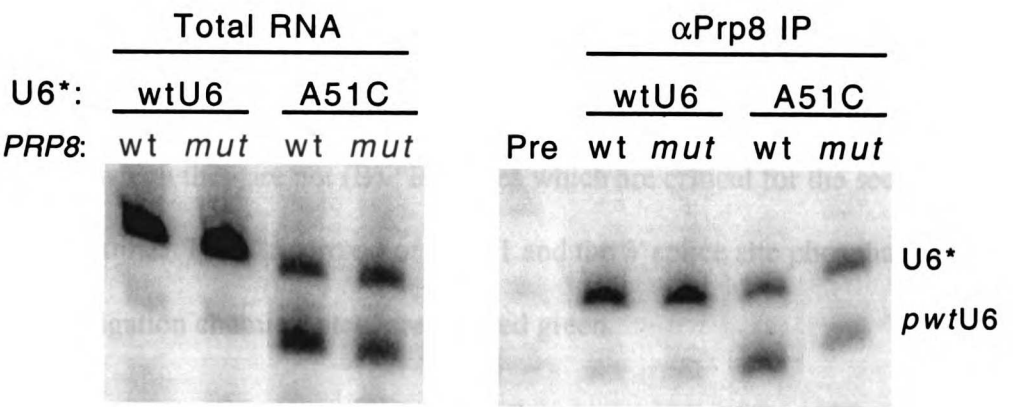
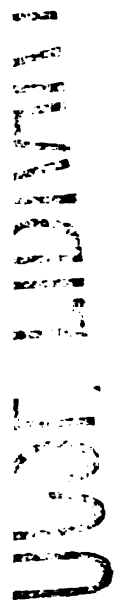
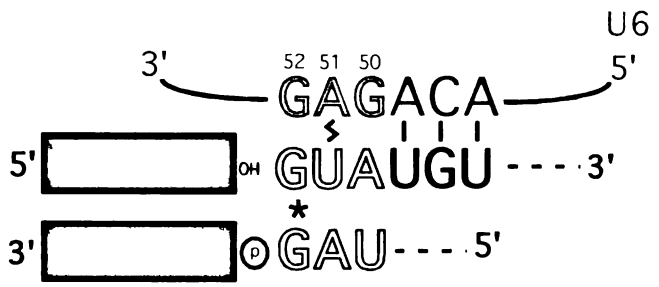


Figure 7: Hypothesized interaction of RNA residues suppressed by *prp8* extends the juxtaposition of conserved residues at the 5' and 3' ends of the intron.

A, B. The previously implicated non-Watson-Crick base-paring interaction between the first and last residue of the intron (*), in addition to Watson Crick (dash) and crosslinking (lightning bolt) interactions between the 5'SS and U6 snRNA, equally favor an alignment of the 5' and 3' ends of the intron in which additional residues are juxtaposed (A), or an alignment in which they are not (B). Residues which are critical for the second step are denoted in outline. The 3' hydroxyl of exon 1 and the 3' splice site phosphate, reactants for the exon ligation chemical step, are colored green.



A


U6
3'
52 51 50
5'
GAGACA
5'
OH
GUAUGU - - - 3'
*
3'
P
GAU - - - 5'

B

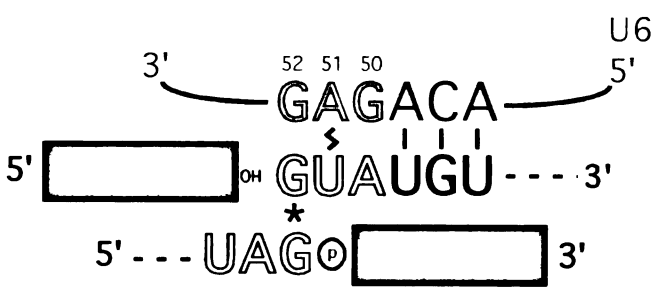
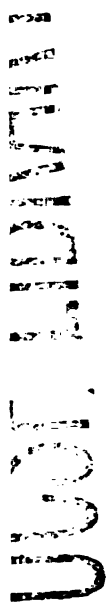
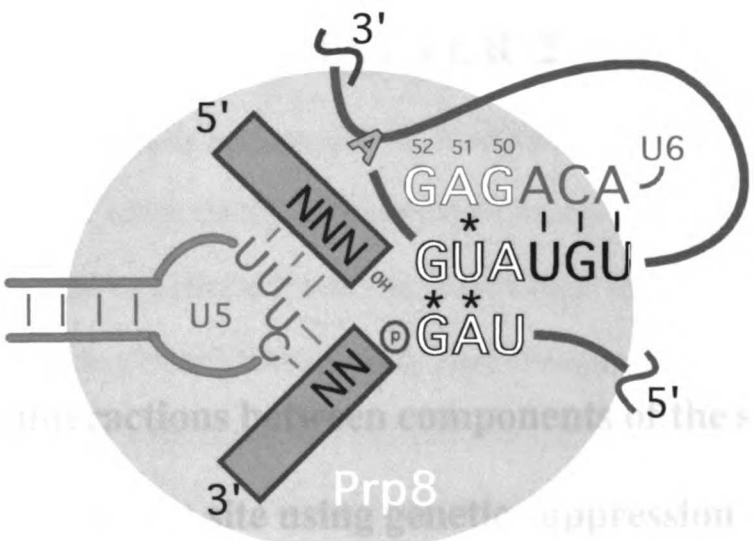

U6
3'
52 51 50
5'
GAGACA
5'
OH
GUAUGU - - - 3'
*
5' - - - UAG
P
3'

Figure 7: Hypothesized interaction of RNA residues suppressed by *prp8* extends the juxtaposition of conserved residues at the 5' and 3' ends of the intron.

C. Prp8 could function in conjunction with RNA interactions at the spliceosomal active site for the second step. The known and hypothesized (asterisks) RNA interactions could align the 5' and 3' ends of the intron, in addition to the 5' and 3' exons, to position the reactants (green) for exon ligation catalysis. Interactions between the invariant loop of U5 snRNA and the 5' and 3' exons are shown in blue.




3'
5'
A
52 51 50
U6
GAGACA
NNN
OH
GUAUGU
U5
UUUC
P
GAU
NN
5'
3'
Prp8

CHAPTER 2

Probing interactions between components of the spliceosomal active site using genetic suppression

ABSTRACT

Previous observations have suggested a model for the catalytic core of the spliceosome, in which consensus sequences near the 5' splice site and the 3' splice site interact with each other during the second step of pre-mRNA splicing. However, only one specific base pairing interaction, between terminal 5' and 3' splice site residues, has thus far been indentified. The results from the previous chapter suggested that the protein Prp8 simultaneously interacts with the 3' splice site, position 2 of the 5' splice site, and position 51 within the highly conserved ACAGAG motif of U6. Thus an extended interaction between the 5' and 3' splice site might involve additional components, including U6 snRNA, and Prp8. Here I describe a genetic screen that provides evidence for a specific tertiary RNA structure, consisting of standard base pairs or triples, forms between 5' and 3' intron residues and/or U6 snRNA. While systematic analysis of combinations of mutations in these residues does not support the existence of standard base pairs or triples, a striking genetic suppression interaction was observed between mutations in position +3 of the 5' splice site, position -3 of the 3' splice site, and position 50 of U6 snRNA. The effects of U6 mutations upon specific **positions** of the 5' and 3' splice sites suggests a specific alignment for the two ends of the intron. These observations, taken together with further characterization of the *prp8* splice site suppressor alleles, provide additional support for the model that Prp8 influences a tertiary RNA interaction at the catalytic core.

INTRODUCTION

A structural model for the spliceosome's catalytic core remains an elusive goal. However, many interactions between RNA components of the core have been inferred through characterization of allele-specific genetic interactions. The previous chapter described a striking array of genetic interactions between a protein component of the catalytic core, Prp8, and specific RNA residues. To explain the ability of the *prp8* 'splice site suppressor' alleles to simultaneously rescue a distinct subset of spliceosomal RNA mutations, a model was proposed in which Prp8 influences a tertiary interaction between critical residues in the catalytic core. Specifically (Figure 1), an interaction between the 5' and 3' intron splice site consensus sequences that includes residues in the U6 ACAGAG motif, and Prp8, would add significant constraints and information to the working model of the spliceosome's catalytic core. (For instance, this interaction would argue against the existence of such configurations as Figure 1B). It would also provide a rationale for the very stringent requirement for all of these highly conserved RNA residues in the second catalytic step of splicing (exon ligation).

This chapter describes genetic analysis of the spliceosome's catalytic core with the purpose of addressing two questions. One is whether there is a direct interaction between the RNA residues themselves. This question was addressed by screening for allele-specific suppressor interactions between residues in the 5' splice site (5'SS), 3' splice site (3'SS), and U6. The corollary is whether the genetic data imply the existence of a specific base pair or base-triple interaction.

Observations of allele-specific genetic suppression can be particularly powerful for constraining molecular models. A Watson-Crick base pairing interaction can be inferred when a mutation in either of the pairing partners results in a functional defect that can be rescued by an additional mutation that restores the ability to form a Watson-Crick base pair. In some cases, non-Watson-Crick base pairing interactions can also be inferred genetically. In a particularly relevant example, the 5' and 3' terminal residues of the intron have been implicated in forming a base pairing interaction. Mutations in the 5'SS /GUA (the slash indicates the 5' splice site) and in the 3' SS UAG/ (here the slash indicates the 3' splice site) can severely inhibit the second step of splicing (Fouser and Friesen, 1987; Fouser and Friesen, 1986). However, a specific combination of mutations in the two sites, 5'SS /GUA → /**a**UA with 3'SS UAG/ → UA**c**/, allows for splicing to proceed (Chanfreau and Jacquier, 1993; Deirdre et al., 1995; Parker and Siliciano, 1993). The high degree of allele-specificity (no other combinations of mutations work) is best explained by a model in which the terminal G residues of the intron splice sites form a specific interaction that can be functionally substituted by an interaction between the A and C mutations. One possible non-Watson-Crick base pair between two G's, which can be structurally mimicked by a reverse Wobble A-C base pair, is shown in Figure 2. While formal proof of the specific mechanism of interaction requires structural analysis, the genetic data strongly imply the existence of a physical interaction.

Previously, such evidence for an interaction between the 5'SS and 3'SS penultimate residues was looked for by combining all possible mutations in 5'SS position +2 and 3'SS position –2 (/GxA—UxG/) (Ruis et al., 1994). All double mutant combinations were even less efficient for splicing than single combinations. In the

context of the proposed interaction, the failure to observe suppression could be due to constraints imposed by additional components of the interaction, U6 and Prp8. An obvious experiment was then to combine these splice site mutations with mutations in U6 and/or Prp8 and repeat the search for combinations of mutations that can restore function. The proposed tertiary interaction could conceivably juxtapose 5'SS position +3 with 3'SS position –3 and U6 G50. These residues are also highly conserved and important for the second catalytic step. Thus, these residues were also included in the analysis.

Since the catalytic core is currently not amenable to structural studies, the kind of genetic analysis initiated here takes us as close as we can get to characterizing the nature of the hypothesized tertiary interaction. Prp8's role could be to stabilize or constrain an interaction that is primarily between RNA components, or Prp8 could contribute essential amino acids, whose interactions inter-digitate with those of RNA. A result that allowed a specific RNA structure to be modeled, by implying the existence of a specific base pair or triple, would favor the first hypothesized role for Prp8, in influencing a structure that is primarily RNA. If specific base pairs or triples cannot be implied, then resolving hypothesized roles and interactions of Prp8 with RNA must await detailed structural analysis.

The genetic results described in this chapter favor the existence of a tertiary interaction that does not consist of standard base pairs or triples. A striking observation of genetic suppression was made between a specific combination of mutations in the 5' splice site, 3' splice site, and U6. This genetic interaction provides further support for the hypothesis that these RNA residues interact during the second catalytic step of

splicing. However, the genetic interaction does not provide enough constraints for molecular modeling.

This chapter also presents some further genetic characterization of the *prp8* splice site suppressor alleles. I repeated the screens of combinations of RNA mutations for allele-specific suppression in the presence of *prp8* splice site suppressor alleles, with the hopes that this background might allow for better detection of RNA-RNA interactions. I found that the *prp8* alleles exacerbate the effects of many RNA mutations. The observation that the suppression of multiple RNA mutations by *prp8* is not simply additive (and are thus synergistic) provides further evidence for the proposed interaction. Interestingly, some *prp8* alleles mimic the suppression described above of a specific combination of mutations in the 5' and 3' splice sites by mutations in U6. This significance and limitations to interpreting these genetic interactions is discussed.

RESULTS

I. GENETIC INTERACTIONS BETWEEN 5'SS, 3'SS AND U6 RNAS

To explore the hypothesis that residues in the 5' splice site, 3' splice site, and U6 interact, perhaps in the form of base pairs or base triples, I screened all possible specific combinations of mutations in these residues for the ability to confer allele-specific genetic suppression. This approach is also known as 'artificial phylogenetic analysis'. Because the residues in Figure 1 are very highly conserved, there are very little

phylogenetic data available. Thus mutations are generated and screened for combinations that can restore function. This approach (using mutations in specific positions of U2 and U6) has convincingly implicated an intriguing tertiary interaction in the spliceosome (Madhani and Guthrie, 1994).

To assay the effects of intron mutations in living cells, all possible single and double mutations between 5'SS position +2 and 3'SS position –2 (/G**x**A—U**x**G/) (Ruis et al., 1994); and 5'SS position +3 and 3'SS position –3 (/GU**x**—**x**AG/) were generated in *ACT1-CUP1* splicing reporters. The effects of double mutations in 5'SS position +2 and 3'SS position –2 (/G**x**A—U**x**G/) have been previously reported. All double mutatnt combinations are less efficient for splicing than single mutants (Ruis et al., 1994). The effects of 5'SS +3 and 3'SS –3 (/GU**x**—**x**AG/) mutations are summarized in Table 1. Most of the double mutations are not significantly worse than the more severe single mutations. However, two double mutations, /GU**c**—**a**AG/, and /GU**c**—**g**AG/, were far more severe than either of the single mutations (shaded in Table 1).

We should note that the fact that U6 snRNA is essential for growth has limited the presented analysis to *viable* mutations in the residues of interest: U6 G52**u,** G50**c** and G50**u** (G52**u,** G50**u** are temperature sensitive). Mutations in U6 A51 are inviable alone (Madhani et al., 1990), and confer dominant negative effects in the presence of a wild type copy of U6 (Chapter 1). However, not shown, when a wild type copy of U6 was present, some of the A51 mutations could weakly suppress some of the 5'SS +2; 3'ss –2 double mutant reporters (not shown).

Suppression of a 5'SS-3'SS double mutant by U6 G50c and U6 G50u

All of the *ACT1-CUP1* reporters were assayed for copper resistance in strains containing wild type or viable mutations in U6 as the sole copy of U6 (Figure 3). Some of the mutant reporters were weakly suppressed (small '+') by the U6 mutations G50**c** and G50**u**. Strikingly, the double-mutant /GU**c**—**a**AG/ reporter was strongly suppressed by both U6 50**c** (20 fold) and 50**u** (30 fold). The large '+' in Figure 3 indicates these observations of strong suppression.

The data in Figure 3 suggest that U6 G50 and 5'SS +3A directly influence each other. First, consider that the mutations G50**c** and G50**u** weakly suppress 5'SS /GU**c**, and worsen 5'SS /GU**u**. These observations alone are not surprising in light of many previous observations which suggest that these 5'SS and U6 residues are indeed close to each other on the spliceosome. These observations include crosslinking interactions between the adjacent 5'SS +2U and U6 51A (Kim and Abelson, 1996; Sontheimer and Steitz, 1993) and genetically implied basepairing interactions between 5'SS +4,U, +5,G, and +6U with U6 49A, 48C, and 47A (Kandels and Séraphin, 1993; Lesser and Guthrie, 1993).

I wanted to determine whether residues near the 3' splice site are also involved in this interaction. The extremely strong suppression of the /GU**c**—**a**AG/ by U6 G50**c** and G50**u** fits this criterion: Some suppression of double mutant reporter might be expected, since the 5'SS /GU**c** mutation is suppressed approximately 2-fold by G50**c** and 2.5-fold by G50**u** (Table 2). However, the 20-fold and 30-fold suppression of the 5'SS, 3'SS double-mutant /GU**c**—**a**AG/ is much greater. This suggests that the 3' splice site plays a role in the 5'SS-U6 interaction.

Specificity in the position of suppression

There is specificity in the 3'SS mutation that confers suppression. The /GU**c**—**g**AG/ mutation is not suppressed by the U6 mutations, and the /GU**c**—**c**AG/ double mutant is not suppressed to any greater extent than the 5'SS +3**c** single mutant. I also tested for specificity in the **position** of the 3'SS mutation. Double mutant reporters containing mutations at 3'SS –1G are not suppressed by U6 G50**c** and G50**u** (data not shown).

Suppression of +3;-3 mutant reporters also appears to be specific to position G50 in U6. The nearby viable mutation in U6, G52**u**, was tested in parallel with all of the reporter mutations. While G52**u** exacerbated the copper resistance of many mutant reporters, no suppression was observed. In contrast, U6 G52**u** was previously observed to confer suppression to 3'SS –1 mutations (Lesser and Guthrie, 1993). The three viable U6 mutations were also assayed with reporters containing all possible combinations of mutations at 5'SS +2 and 3'SS –2. No suppression was observed (not shown). In contrast, U6 A51**g** could dominantly confer weak suppression to a number of the +2-2 double mutant reporters (not shown). Thus, suppression is very specific to position, and all of the results are consistent with an alignment of the 3'SS with the 5'SS and U6, as drawn in Figure 1. This alignment is now highly favored over the alignment in Figure 1B.

The efficiency of the second step, at the correct splice sites, is increased

I confirmed that the strong increase in copper resistance conferred by the /GU**c**—**a**AG/ reporter with U6 50**c** and 50**u** was indeed due to increased splicing at the correct (mutated) splice sites. The splicing of the of the *ACT1-CUP1* reporters in strains containing wild-type or mutant U6 was assayed by primer extension, using a ^{32}P-labelled oligo (CUP1-CC4) that hyridizes to the second exon of the reporter. Shown in Figure 4, in the presence of wild type U6, the /GU**c**—**a**AG/ reporter generated a cDNA species about 19 nucleotides shorter than the cDNA for the correctly spliced mRNA product. A product of the same size was observed for some 3' splice site mutations (not shown), consistent with the possibility that it is generated by usage of a CAG sequence 19 nucleotides downstream of the correct (mutated) 3' splice site. Usage of this cryptic site would create a frame-shift in the *CUP1* open reading frame, consistent with the low copper resistance of the reporter. U6 G50**c** and G50**u** confer an increase in the level of the correctly spliced mRNA product. Thus suppression occurs through increased splicing at the mutated splice sites.

Using the assumption that the splicing reaction follows first-order kinetics, one can estimate the efficiency of the splicing reaction by the steady-state ratio of Mature/Precursor levels (Pikielny and Rosbash, 1985), Likewise, using the assumption that the second catalytic step, exon ligation, follows first-order kinetics, one can estimate its efficiency with the ratio of Mature/Lariat intermediate levels (Fouser and Friesen, 1986). Levels for each reporter species were measured by PhosphorImager analysis of the primer extension products. Both 5'SS /GU**c** and the /GU**c**—**a**AG/ mutant reporters are strongly impaired for the exon ligation step. U6 50**c** and U6 50**u** increase the efficiency

of exon ligation for these mutant reporters. (Figure 4 and Table 2). For the /GUc—aAG/ double mutant, this increase is particularly dramatic, possibly due to additive effects of increasing the efficiency of ligation at the mutant splice sites and decreasing the usage of the downstream cryptic splice site.

Thus, the highly conserved residues, 5'SS +3A, and U6 G50 appear to influence the identity of the nucleotide at 3'SS –3. Mutations in all of these residues affect the second catalytic step, exon ligation, and the data above are consistent with the idea that their critical function at the second step is mediated by a tertiary interaction between them.

II. EFFECTS OF *PRP8* MUTATIONS

The previous observation that alleles of *prp8* can permit the splicing of mutations in the 5' splice site, 3' splice site, and U6 (Chapter 1) suggests that Prp8 may function to constrain the hypothesized tertiary interaction. I was thus curious to investigate the effects of combining mutations in the 5' splice site, 3' splice site, and U6, in a *prp8* background. Perhaps the *prp8* mutation, by imposing weaker constraints upon the catalytic core, would allow for better detection of suppressor interactions between mutations in RNA. Alternatively, Prp8 might add stability to the interaction, which could become important when additional components of the catalytic core are mutated.

Non-additive effects of prp8 upon 5'SS +2 and 5'SS –2 mutations

As was done for the viable U6 mutations, I assayed the effects of two *prp8* suppressor alleles, *prp8-C122* and *prp8-D136*, upon all of the *ACT1-CUP1* reporters containing double mutations in +2,-2 or +3-3 (using the copper resistance assay).

Notably, when mutations that alone were suppressed by the *prp8* alleles (Chapter 1), for example 5'SS /G**a**A, and 3'SS U**g**G/, were combined, the resulting double mutant reporter was not suppressed (Figure 5). If the suppression by *prp8* at the two distinct splice sites occurs independently, then it should be additive (the total suppression would be the product of the extent of suppression at each site). The effects of the +2-2 double mutations are <u>less</u> than the effects upon any single +2 or –2 mutation, consistent with effects of Prp8 at the two sites being interrelated. Probably, suppression requires some approximation of a tertiary interaction between the two sites, which the double mutation might abolish completely.

Some prp8 alleles mimic U6 G50 mutations by suppressing a +3-3 double mutant reporter

Out of all +3-3 double mutant reporters tested with the two *prp8* alleles, only one effect was observed: *prp8-C122* suppresses /GU**c**—**a**AG/, conferring an approximately 10 fold increase in copper resistance (Figure 6 and data not shown). Strikingly, this is the same double mutant reporter that is strongly suppressed (20-30 fold) by U6 50**c** and U6 50**u**.

Intrigued by the observation that only one of the *prp8* alleles tested suppressed this reporter, I proceeded to assay more of the *prp8* alleles from my collection for effects upon the /GUc—aAG/ double mutant *ACT1-CUP1* reporter and the single mutants 5'SS /GUc, and 3'SS aAG/. A number of the 'splice site suppressor' alleles (Chapter 1) of *prp8* indeed suppress the /GUc—aAG/ reporter, and a number do not (Figure 6 and data not shown). A few of the *prp8* alleles exacerbate the 5'SS /GUc single mutant reporter.

There is no apparent correlation between suppression of /GUc—aAG/ and any other observed property of the *prp8* alleles. These properties include (1) the effects observed upon the 5'SS /GUc single mutant reporter. (One allele, *prp8-C121*, exacerbates 5'SS /GUc while suppressing /GUc—aAG/). (2) Alleles that suppress /GUc—aAG/ do not map to any distinct regions of the protein. Some alleles that arise from mutations in the B, C, or D regions of the protein can confer suppression. (3) In Chapter 1, some of the splice site suppressor alleles of *prp8* did not confer very strong suppression of U6 A51 mutations (Ch.1 Figure 6 and data not shown). This property does not correlate with /GUc—aAG/ suppression. (4) Some of the splice site suppressor alleles confer cold-sensitive growth (not shown). This phenotype also does not correlate with /GUc—aAG/ suppression.

Could some of the *prp8* alleles display genetic interactions with U6 G50 mutants? Described in Chapter 1, a few *prp8* alleles were combined with mutations in U6 G50, and no effect was observed. However, I did not test the entire array of *prp8* alleles.

Quadruple mutants

To combine *prp8* mutants with triple RNA mutations in the 5'SS, 3'SS, and U6, I needed to construct strains in which the *prp8* alleles were chromosomally integrated. I used the pop-in/pop-out allele replacement strategy (Rothstein, 1991) for recombining the mutant allele into the genome and screened for successful substitution of the *prp8* allele by assaying for previously observed splice site suppression. Unfortunately, even when wild type *prp8* was used with this strategy, splice site suppressors were isolated. A possible explanation is that this strategy generates new *prp8* splice site suppressor alleles through mutagenesis during recombination. This needs to be formally tested by sequencing *PRP8* in these strains. For now, I will report effects in these strains, $prp8^1$ and $prp8^2$, upon combinations of RNA mutations, with the caveat that I do not know exactly the *prp8* mutation tested.

For simplicity, I examined the case in which the /GUc—aAG/ double mutant reporter is suppressed by mutations U6 G50c and G50u. Two isolates from the *prp8* integration exhibit slightly different effects.

One isolate *($prp8^1$)* suppresses the /GU**c**—**a**AG/ double mutant reporter, thereby somewhat mimicking the suppression by U6 G50**c** and G50**u**. The suppression is not as strong as that by the U6 mutants, and interestingly, not additive with suppression by U6 G50**c** and 50**u**. That is, the quadruple mutant is slightly worse, not better than, the /GU**c**—**a**AG/; U6 G50**c** or 50**u** triple mutants. (Table 3). This non-additive effect is again consistent with an interrelated function for each of the RNA residues and Prp8.

The other isolate *($prp8^2$)* does not suppress the /GU**c**—**a**AG/ double-mutant reporter and exacerbates the 5'SS /GU**c** single mutant reporter. Despite the effect upon

the single mutant reporter, this mutant does not interfere with or enhance the suppression by mutations in U6 snRNA.

DISCUSSION

This series of genetic experiments was undertaken with the goal of determining whether critical RNA residues in the catalytic core, which have previously implicated in interacting with the protein Prp8, indeed form a direct tertiary interaction with each other. A corollary question was whether the genetic data we obtained could imply the existence of a specific base pair or base-triple interaction. The existence of such an interaction would provide further constraints for modeling the structure of the catalytic core and would provide further evidence that role of Prp8 is to stabilize or constrain a structure whose fundamental interactions are comprised of RNA.

Indeed, I found further genetic evidence for a tertiary interaction between the 5' splice site, 3' splice site, and U6 snRNA, supporting the model proposed in Figure 1A. A number of suppressor interactions were observed between mutations in U6 and the 5' splice site. Some of these interactions were much stronger when combined with specific mutations in 3' splice site. Strikingly, all the observations suggested an alignment and ***register*** of the residues as drawn in Figure 1A. Mutations in U6 G50 could only suppress mutations in 5'SS +3A or 3'SS –3U (Figure 3). Likewise, mutations in U6 A51 only affected mutations in 5'SS +2U or 3'SS –2A (not shown).

Base triples or base pairs?

In an attempt to interpret the genetic results with possible base paring or base-triple interactions that could form between the residues of interest, I made use of a searchable database of base pair and base triple interactions, developed by Bernhard Walberer and Alan Frankel (Walberer, 2000). This database was constructed by enumerating all sterically feasible orientations of the bases (in one plane) that allow for formation of at least two hydrogen bonds between each base. Structural properties of the base interactions, such as the relative locations of glycosidic bond linkages, are annotated so that they can be easily compared to one another. By 'mining' this database using the program MINE (Walberer, 2000), base pairs or triples that are structurally similar (display structural 'overlap'), can be identified.

I focussed on the most striking genetic interaction: the observation that the /GU**c**-**a**AG/ double mutant reporter was rescued to near wild type levels by mutations in U6, G50**c**, and G50**u**. Representing the components of the wild type spliceosome, the MINE database contains 35 possible $G_{(U6\ 50)}$ - $A_{(5'SS\ +3)}$ – $U_{(3'SS-3)}$ base triples. However, none of these show much isostery with any of the possible triples formed by the combinations of mutations that confer suppression: C-C-A or U-C-A.

Because of the extensive evidence that the 5' splice site and U6 do interact, I also considered the possibility that a base pair forms between 5'SS +3A and U6 G50. Of the four possible G-A basepairs, one has modest structural overlap with a possible C-U pairing, suggested by the suppression of 5'SS +3**c** by U6 G50**u** (Figure 7). Since this G-A base pair also has strong structural overlap with Watson-Crick base pairs, a prediction of this model is that combining mutations that allow formation of a Watson-Crick base

pair between these positions might also restore function. However, this prediction is not supported by my data. Notably, I found that any mutations that allowed Watson-Crick pairing between any two positions of 5'SS +3, 3'SS –3 and U6 G50 were deleterious (Figure 3).

Thus, the genetic suppression data in hand provide no evidence for the existence of standard base pairs or triples. Importantly, the base triple interactions that were evaluated by modeling cannot exhaustively represent all the possible ways in which three RNA residues can interact. Many base triples that have been observed in nature involve only a single hydrogen bond from one of the bases. Bases do not always interact in one plane, and some triple interactions have been observed to involve contributions from the RNA-phosphate backbone and from bound water molecules (for example, Boudvillain et al., 2000; Szewczak et al., 1998). It is not computationally possible for the database of triples to include and evaluate all of these possibilities. Given the many forms of RNA interactions that have now been identified through structural analysis, it may not be surprising (in retrospect) that this analysis did not yield a simple base pairing or base triple model.

Another limitation to the modeling attempts is that there is no standardized method for evaluating the genetic data. Because no example of suppression was observed to restore function to 100% of the wild-type level, it was difficult to evaluate how much structural difference could be reasonably tolerated between modeled interactions involving wild type or mutated residues. My only recourse in light of this issue was to consider only the strongest suppressor interactions, in which splicing was restored close to wild-type levels.

I also attempted to model the proposed interaction using some artificial constraints instead of the suppression data. For instance, I considered the assumption that the interaction forms a standard triple helix, which requires each base triple to stack upon the adjacent triple with the glycosidic bonds in similar orientations. The MINE database was used to evaluate various stacking configurations of the three hypothesized triples: $G_{(U6\ 50)}$ - $A_{(5'SS\ +3)}$ – $U_{(3'SS-3}$; $A_{(U6\ 51)}$ - $U_{(5'SS\ +2)}$ – $A_{(3'SS-2;}$; and $G_{(U6\ 52)}$ - $G_{(5'SS\ +1)}$ – $G_{(3'SS-1)}$. There were no sets of triples containing these sequences within the MINE database that satisfied any of the constraints or assumptions for a standard triple helix.

Allele specificity

Evaluation of specific models for the proposed tertiary interaction is also limited by the fact that the genetic suppression observed here is only modestly allele-specific. Both U6 G50c and G50u suppress the splicing of a few *ACT1-CUP1* mutant reporters. Notably, suppression that is strongly specific for the *position* of the mutation, but not so specific for the *identity* of the mutation, is reminiscent of other genetically inferred spliceosomal interactions. For instance, the mutation in U6 G52u allows for suppression of several mutations in the last residue of the 3'SS UAG/ (Lesser and Guthrie, 1993). Also several mutations near the bulge of U2/U6 helix I allow for suppression of lethal mutations in U6 G52 (Madhani and Guthrie, 1994). Lastly, the *prp8* alleles described in Chapter 1 allow for several mutations at specific positions in the 5'SS, 3'SS and U6 to be tolerated. While a physical interaction is a likely explanation for suppression in all of these instances, a specific kind of interaction (such as base pairing) cannot be implied.

Formally, the suppressed residues could influence each other indirectly, via one or more additional factors.

In the first Chapter, I discussed how suppression of splice site mutations by *prp8* alleles might occur through a loosening of constraints upon specific spliceosomal residues. Such effects could be elicited by an indirect interaction between the RNA and the mutated Prp8 residues. Suppression by U6 G50c and U6 G50u might occur by a similar mechanism. Consistent with this idea, both U6 G50 mutants and some *prp8* alleles suppress the same double mutant intron reporter, /GUc—aAG/.

However, if suppression occurs by a loosening of constraints, the effect is not entirely general. Suppression of intron mutations by U6 mutantions is specific to position. Also, with respect to the entire array of different splice site mutations, the individual effects of each *prp8* allele, and the U6 alleles, are unique. For instance, many *prp8* splice site suppressor alleles do not suppress the /GUc—aAG/ double mutant reporter (Figure 6). Some *prp8* alleles exacerbate the 5'SS /GUc mutant reporter (Figure 6), while U6 G50c and U6 G50u confer weak suppression (Figure 3). The 5'SS /Guu reporter is exacerbated by U6 U6 G50c and U6 G50u, but not by the *prp8* alleles. (Numerous other examples exist). Of course, no intron mutation is suppressed by all of the alleles tested. One might imagine that the distinct effects upon specific intron positions by different spliceosomal mutations reflect what is possible and not possible in the context of a specific tertiary structure.

Higher resolution suppression experiments using nucleotide analogs

The *ACT1-CUP1* splicing reporter has provided a valuable assay for assessing the effect of intron mutations *in vivo* and for identifying suppressor interactions. However, the kinds of suppressor interactions that can be observed are limited by the possible mutations that can be made: RNA synthesized by only four possible ribonucleosides. In contrast, RNA with a wide variety of modifications, such are removal of 2-hydroxyl group from a sugar at a specific position or removal of the primary amine from a guanosine, can be synthesized using *in vitro* transcription reactions, direct oligonucleotide chemical synthesis, or a combination of these techniques. The powerful technique of nucleotide analog interference suppression (NAIS) makes use of such synthetic RNA technology and has yielded evidence for specific tertiary interactions within several self-splicing RNA (Strobel, 1999). Once a site-specific modification that inhibits splicing has been identified, one can screen through libraries of different modified RNAs (representing all the possible positions) for second-site modifications that can rescue the splicing reaction. In theory, this technique could be applied to the spliceosome, using an *in vitro* splicing reaction. While a synthetic pre-mRNA substrate could be modified easily, synthetic modified snRNAs would need to be reconstituted into snRNPs. Based on other studies with multi-step reactions, an additional challenge will probably lie in finding conditions for which the splicing assay reflects effects at the specific step during which the interaction is functionally important (Gordon et al., 2000; Sontheimer et al., 1999).

MATERIALS AND METHODS

Strains and reagents

All +3, -3 single and double mutant *ACT1-CUP1* reporters (pCC168 through pCC 183, and pCC188 through pCC191) were constructed in the pGAC14 (*TRP*) vector by Quick-Change mutagenesis (Stratagene). For all of these reporters, position 305 has also been mutated to eliminate a cryptic AG 3' splice site: (...UAG/AG to UAG/A**u**). Each construct was sequenced to confirm identity of the introduced mutations. A few double mutant reporters were also subcloned into the pGAC24 (LEU) backbone (pCC184-186).

The +2, -2 double mutant *ACT1-CUP1* reporters, described originally in (Ruis et al., 1994), were provided by Paul Siliciano.

U6 mutants (pCC134-137; 145, 146, and108) were subcloned pSE362 (CEN, HIS) from constructs described in (Hiten and Remy).

prp8 alleles were carried on pRS424 (2 micron, TRP) plasmids (Chapter 1), or pRS313 (CEN, HIS). (pCC197, pCC198, and pCC125).

The *cup1* deletion strains are derived from the strains described in (Lesser and Guthrie, 1993). U6 mutants were assayed in the double deletion strain yCC30: *cup1Δ::ura3, U6Δ::LEU2; his3, ura3, leu2, trp1; WT U6-URA*, which was constructed by Cammie Lesser, probably by crossing yCL42 with yHM1. U6 WT or mutant plasmids (in pSE362) were co-transformed into this strain with each *ACT1-CUP1* reporter. The WT U6-URA plasmid was then shuffled out by growth on 5-FOA at room temperature. The effect of *prp8* mutants were assayed in the *prp8Δ, cup1Δ* double

deletion strain, yJU75, described in (Umen and Guthrie, 1996). The WT PRP8-URA plasmid in this strain was replaced with pRS313 or pRS424 plasmids containing WT or mutant *prp8* alleles by plasmid shuffle (on 5-FOA).

Copper growth, Primer Extension

Copper resistance and primer extension assays, to measure the splicing efficiency of the *ACT1-CUP1* reporters, were conducted as previously described in Chapter 1.

TABLE 1

Copper resistance of 5'SS position +3 and 3'SS position –3 *ACT1-CUP1* reporters

		3'SS			
		UAG/	cAG/	aAG/	gAG/
5'SS	/GUA	2.0	2.0	1.75	0.1
	/GUc	0.3	0.3	0.025	0.025
	/GUg	0.2	0.1	0.2	0.1
	/GUu	0.5	0.2	0.2	0.2

Shaded are double mutants that are significantly worse than either of the respective single mutants.

TABLE 2

Fold suppression of splice site mutations by mutations in U6

/5'SS — 3'SS/		(a) Copper Growth	(a) Fold increase	(b) [Mature]/[Precursor]	(b) Fold increase	(c) [Mature]/[Lariat Int.]	(c) Fold increase
/GUA—UAG/							
	WT U6	2.0	-	13 *+/- 1.7*	-	122 *+/- 46*	-
	G50c	1.75	0.9x	7.3 *+/- 0.8*	0.6x	43 *+/- 5.0*	N/D
	G50u	1.5	0.75x	5.6 *+/- 0.5*	0.4x	67 *+/- 0.5*	N/D
/GUc--UAG/							
	WT U6	0.3	-	1.3 *+/- 0.2*	-	4.7 *+/- 1.0*	-
	G50c	0.5	**1.7x**	1.8 *+/- 0.3*	**1.4x**	37 *+/- 12*	**8x**
	G50u	0.75	**2.5x**	1.7 *+/- 0.3*	**1.3x**	55 *+/- 11*	**12x**
/GUA--aAG/							
	WT U6	1.75	-	10.7 *+/- 0.8*	-	31.6 *+/- 6.7*	-
	G50c	1.5	0.9x	6.3 *+/- 0.05*	0.6x	9.3 *+/- 0.8*	0.3x
	G50u	1.5	0.9x	5.1 *+/- 0.3*	0.5x	39.0 *+/- 4.0*	1.2x
/GUc--aAG/							
	WT U6	0.025	-	0.2 *+/- 0.02*	-	0.06 *+/- 0.005*	-
	G50c	0.5	**20x**	2.2 *+/- 0.1*	**11x**	4.5 *+/- 0.3*	**75x**
	G50u	0.75	**30x**	3.9 *+/- 0.4*	**20x**	12.8 *+/- 0.8*	**213x**

The level of copper resistance (mM copper tolerated for growth) for each strain was measured (three independent times) for (a). For (b) the total splicing efficiency is estimated by the ratio of steady state levels of precursor to lariat intermediate species, measured by phosphorimager analysis. These numbers are the average of two to six independent primer extension experiments. For (c), the efficiency of the second catalytic step of splicing as estimated by the ratio of measured levels of mature to lariat intermediate. (N/D indicates estimates that would be meaningless, given the large error for the measurement in the presence of WT U6). Estimates for cases in which U6 G50c and G50u confer an increase in the splicing efficiency (compared to WT U6) are indicated in bold. The fold suppression by U6 G50c and G50u is much greater for the double mutant reporter /GUc--aAG/ than either single mutant reporter.

TABLE 3

Effects of *prp8 mutants upon triple mutations in the 5'SS, 3'SS and U6 RNAs**

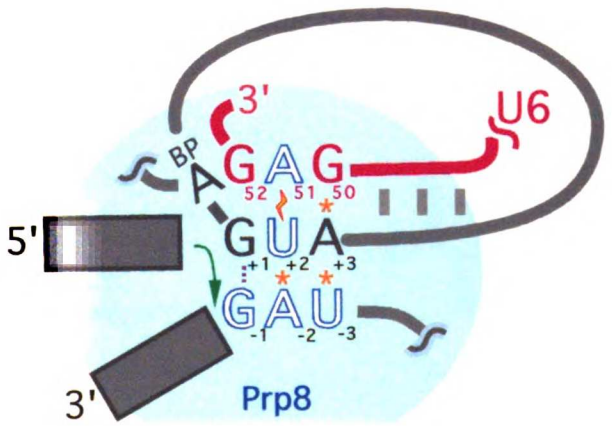
	WT U6	U6 50c	U6 50u
/GUc--UAG/			
WT PRP8	0.3	0.5	0.75
prp8^{1}	0.3	N. D.	N. D.
prp8^{2}	0.1	0.5	0.75
/GUc--aAG/			
WT PRP8	0.025	0.5	0.75
prp8^{1}	0.5	0.5	0.5
prp8^{2}	0.025	0.5	0.5

Numbers indicate the concentration of copper tolerated for growth (mM) by each mutant strain. In outline is the sole example of a *prp8** mutant suppressing a double-mutant *ACT1-CUP1* reporter, and shaded are two instances in which the *prp8** mutants worsen the copper resistance of mutant reporters. N. D. means not determined. *Mutations in *prp8* have not yet been mapped and verified in theses mutant strains.

Figure 1: Model for a tertiary interaction between the 5' SS (+1, +2, +3), 3'SS (-1, -2, -3), U6 (50, 51, 52), and Prp8 at the second catalytic step of splicing.

Both Model **A** and **B** are consistent with previously observed crosslinking interactions (Kim and Abelson, 1996; Sontheimer and Steitz, 1993), indicated by the lightning bolt, and a genetic suppression interaction (Parker and Siliciano, 1993), indicated by the dotted lines. However, Model A. is favored by the genetic interactions of *prp8* described in Chapter 1. In outline are residues that, when mutated, are suppressed by alleles of *prp8*. Asterisks indicate additional proposed interactions, which are supported by observations reported in this Chapter. BP indicates the branchpoint adenosine.

A


3'
U6
BP
A
G A G
52 51 50
5'
G U A
+1 +2 +3
G A U
-1 -2 -3
3'
Prp8

B

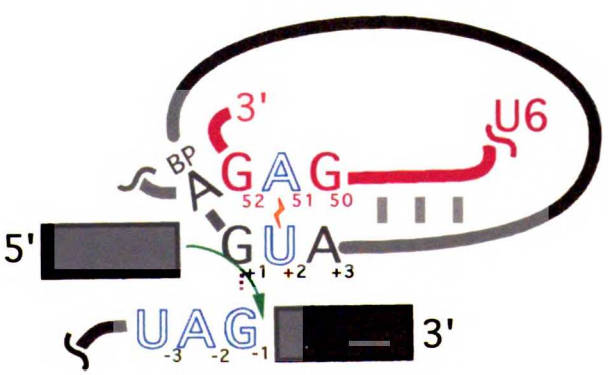
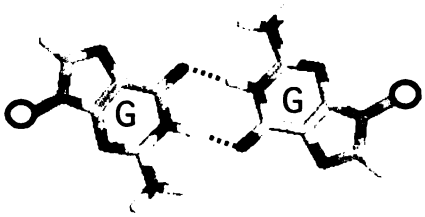

3'
U6
BP
A
G A G
52 51 50
5'
G U A
+1 +2 +3
U A G
-3 -2 -1
3'

Figure 2: Isosteric non-Watson Crick basepairs.

The ability of an A-C double mutation to substitute for the terminal Gs (and the ability of inosines to substitute both Gs (Deirdre et al., 1995)) of the intron could be explained by the formation of a G-G basepair that can be structurally mimicked by a reverse Hoogsteen A-C basepair. Note the similar locations of glycosidic bonds (open circles).


G
G

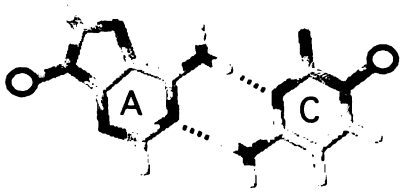

A
C

Figure 3: Copper growth of combinations of mutations in 5'SS position +3, 3'SS position -3, and viable mutations in U6.

Plus symbols indicate mutant *ACT1-CUP1* reporters that are suppressed by U5 G50c and/or G50u.

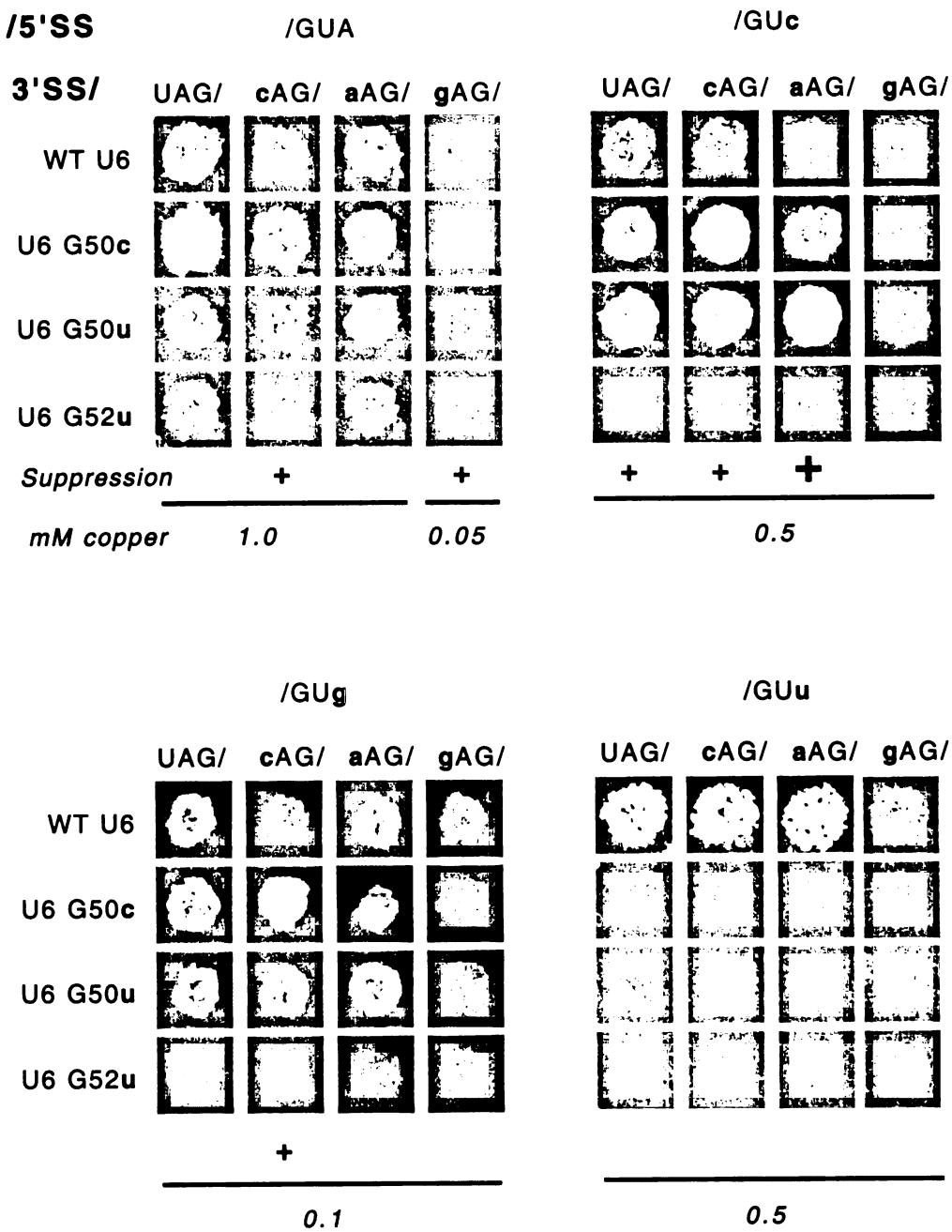

/5'SS
3'SS/
/GUA
/GUc
/GUg
/GUu
UAG/ cAG/ aAG/ gAG/
WT U6
U6 G50c
U6 G50u
U6 G52u
Suppression
mM copper
1.0
0.05
0.5
0.1
0.5

Figure 4: Primer extension data suggest that that the second step splicing defect is suppressed.

Bands corresponding to mature, lariat intermediate, and precursor species are indicated in cartoon on the side. The asterisk indicates species likely to be derived from usage of a downstream CAG sequence as a cryptic 3' splice site. The lariat intermediate species do not line up exactly because the samples were run on two separate gels. Cases in which U6 mutations increase the splicing efficiency are indicated by + symbols.

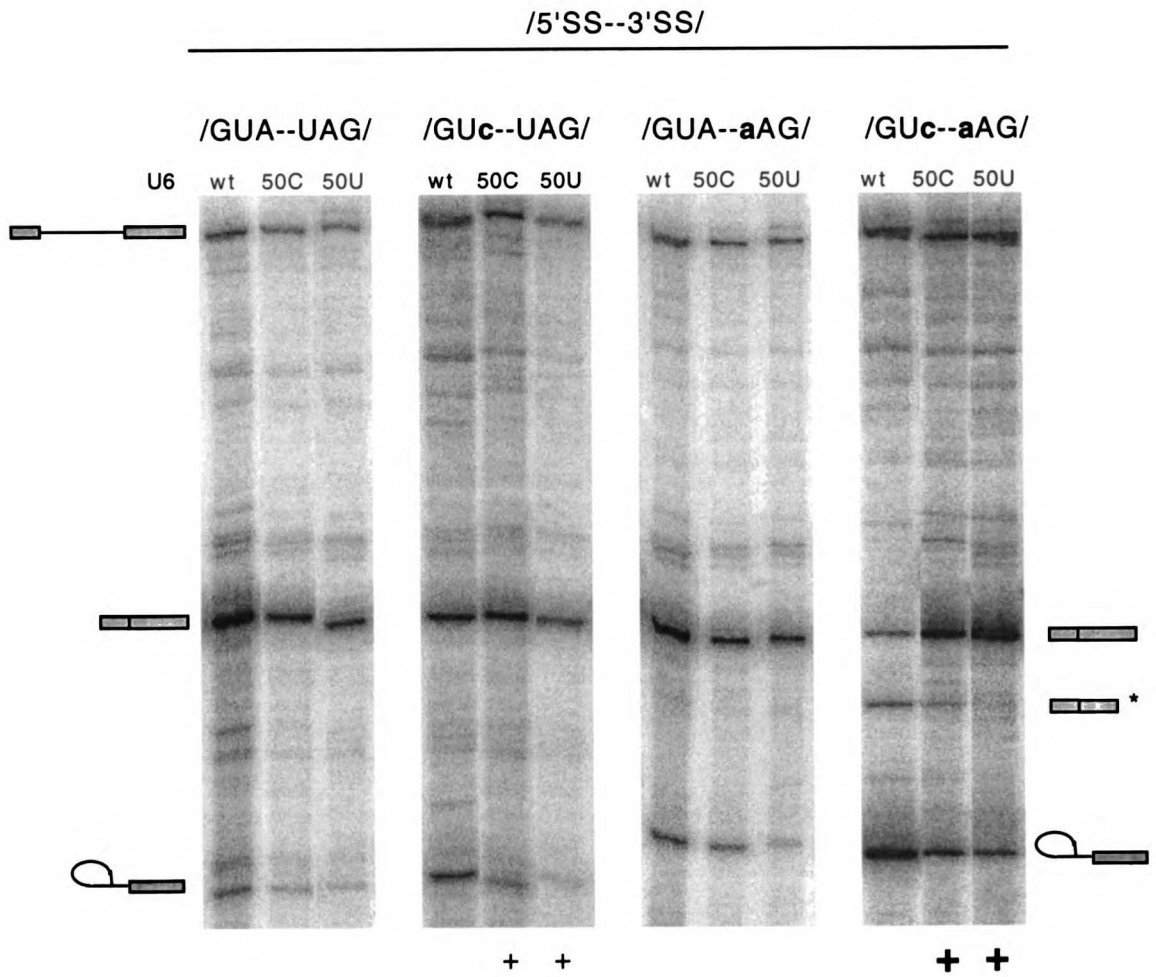

/5'SS--3'SS/
/GUA--UAG/
/GUc--UAG/
/GUA--aAG/
/GUc--aAG/
U6
wt 50C 50U
wt 50C 50U
wt 50C 50U
wt 50C 50U
+ +
+ +
*

Figure 5: Suppression of 5'SS +2 and 3'SS -2 mutations by *prp8* alleles is not additive. The *Δcup1, Δprp8* strain yJU75 contains the indicated *prp8* allele as the sole, plasmid-borne (pSE362) copy of *PRP8*, and a TRP-marked *ACT1-CUP1* reporter plasmid.

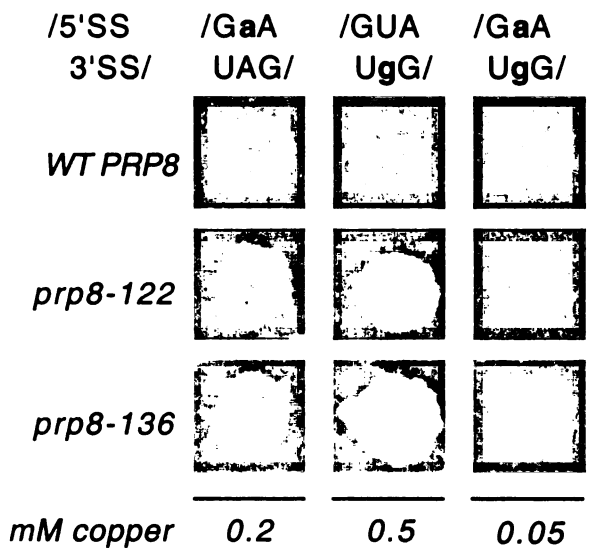

/5'SS
3'SS/
/GaA
UAG/
/GUA
UgG/
/GaA
UgG/
WT PRP8
prp8-122
prp8-136
mM copper
0.2
0.5
0.05

Figure 6: Some, but not all, *prp8* splice site suppressor alleles suppress a 5'SS +3, 3'SS -3 double mutant reporter.

The first column shows copper resistance conferred by different prp8 alleles upon the /GU**c**—**a**AG/ *ACT1-CUP1* reporter. The second column shows that some prp8 alleles exacerbate the copper resistance of the /GU**c**—UAG/ single mutant reporter. The Dcup1, Dprp8 strain yJU75 contains the indicated *prp8* allele as the sole, plasmid-borne (pRS424) copy of *PRP8*, and a LEU-marked *ACT1-CUP1* reporter plasmid.

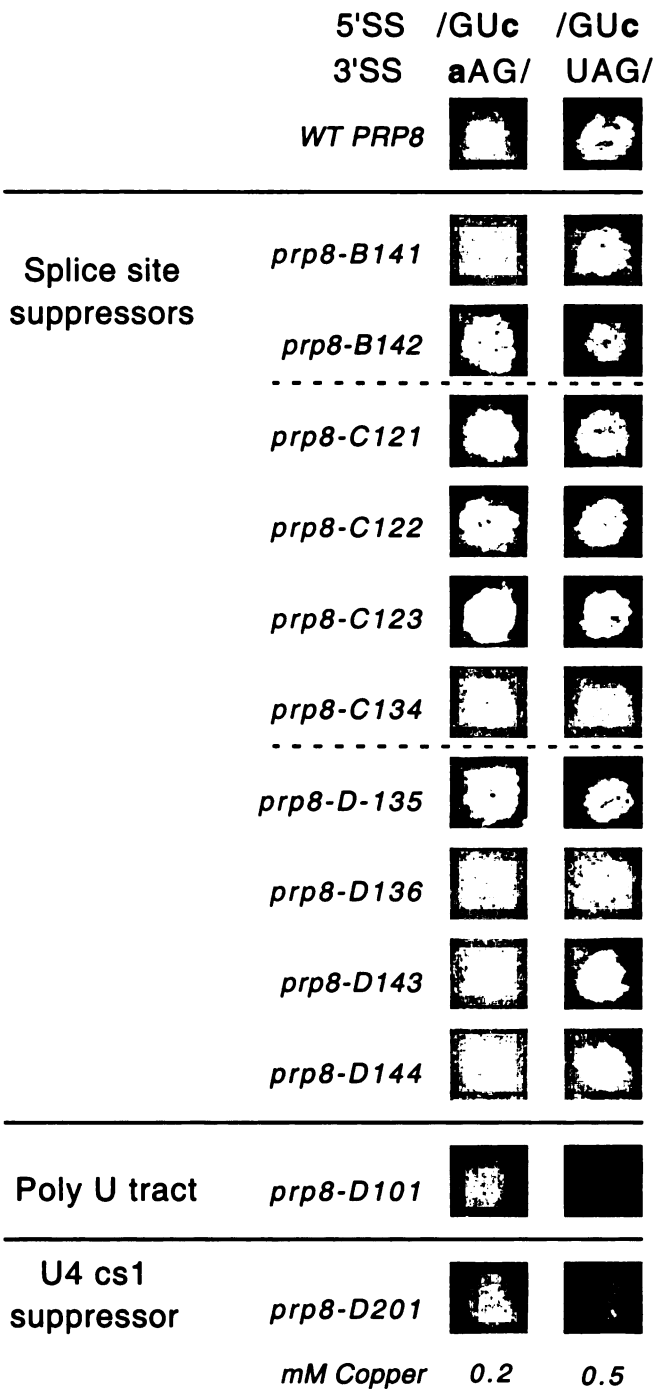
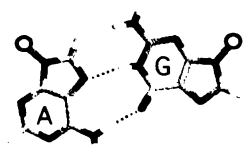
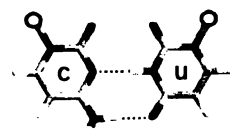

5'SS /GUc /GUc
3'SS aAG/ UAG/
WT PRP8
Splice site suppressors
prp8-B141
prp8-B142
prp8-C121
prp8-C122
prp8-C123
prp8-C134
prp8-D-135
prp8-D136
prp8-D143
prp8-D144
Poly U tract
prp8-D101
U4 cs1 suppressor
prp8-D201
mM Copper 0.2 0.5

Figure 7: Possible non-Watson-Crick basepair between 5'SS +3A and U6 G50. This base pair, with the data-base ID of AG29, displays modest structural overlap (note location of glycosidic linkages, open circles) with a base pair that could explain the suppression of 5'SS +3c by U6 G50u. These base pairs also display strong structural overlap with Watson-Crick base pairs.

5'SS +3 U6 50



AG 29



CU 253

CG 160

CHAPTER 3

An assay for identifying RNA contacts to the 3' splice site UAG

INTRODUCTION

Intron consensus sequences at the 5' splice site, branch point, and 3' splice site play important roles in determining the sites of spliceosome assembly and splicing (reviewed in (Burge et al., 1998)). At least some of the RNA residues in these sequences also appear to play critical roles in catalysis. The 5' (GUA) and 3' (UAG) residues immediately adjacent to the two sites of splicing are required for the second step. The previous two chapters discuss the hypothesis that these residues are critical components of the spliceosomal active site, or catalytic core, that carries out catalysis of the phosphoryl transfer reaction.

Previous crosslinking results provide evidence for physical contacts between the residues in the 5' splice site GUA with U6 (Kim and Abelson, 1996; Sontheimer and Steitz, 1993) and Prp8 (Reyes et al., 1996). However, RNA contacts made by the 3' splice site UAG have never been reported. The specific model put forth in Chapter 1 predicts that these residues contact the 5' GUA and/or U6 residues in the ACAGAG motif. Previously observed genetic interactions also implicate residues in U2/U6 helix I as candidates for interaction with the 3' splice site UAG (Chang and McPheeters, 2000; Madhani and Guthrie, 1994).

To further characterize the RNA-RNA interactions of the catalytic core, I have developed a crosslinking assay for identifying physical contacts of the 3' splice site UAG. The assay makes use of available technology for introducing a photo-activatable nucleotide analogue at a specific position within a synthetic pre-mRNA substrate for *in vitro* splicing. 4-thio-uridine (4-S-U) and 5-iodo-uridine (5-I-U) are two readily

available nucleotide analogues that can be specifically activated by low wavelength UV light to form photo-adducts to other nucleotides and amino acids that lie within a few angstroms (Favre et al., 1998; Hanna, 1989). These analogues were incorporated specifically at position –3 of the 3' splice site UAG in a synthetic RNA substrate for *in vitro* splicing. In the course of an *in vitro* splicing reaction, the 4-S-U substrate forms at least two specific RNA crosslinks. Preliminarily, one of the crosslinks appears to be to U2 snRNA.

RESULTS

Generation of a 3' SS photocrosslinking substrate

The strategy for synthesizing the crosslinking substrate was modified from an original design by David McPheeters (McPheeters et al., 2000). Outlined in Figure 1, it is constructed by ligation of three RNA pieces. The 5' and 3' pieces are synthesized through *in vitro* transcription reactions, and a 13-mer RNA containing the modified (or, for control, unmodified) uridine was chemically sythesized by Dharmacon (Boulder, CO). A long DNA oligo complementary to the RNA oligo and to 20 nucleotides of the 5' piece, and 16 nucleotides of the 3' piece was annealed with all three RNA pieces to form a substrate for T4 DNA ligase. The 5' end of the 3' piece was treated with phosphatase, then kinased with γ^{32}P-ATP. The 4-S-U and 5-I-U modifications did not inhibit the splicing efficiency of the substrate (not shown).

Crosslinking

The substrate was incubated with wild-type yeast nuclear extract under splicing conditions for 15 minutes, then irradiated (at 4°C) with either a 316 nm or 366 nm UV lamp. After extraction and electrophoresis on a denaturing gel, RNA-RNA crosslinked species should be visible by autoradiography as aberrantly (super-shifted) migrating bands. The 4-thio-uridine adduct formed many such species when irradiated with a 366 nm lamp (Figure 1). The 5-iodo-uridine modification may require a higher energy than provided by our 316 nm lamp in order to activate efficient RNA crosslinking (D. Ryan and J. Abelson, personal communication).

In order to detect crosslinks, the reactions were performed on a relatively large scale (6x more extract and substrate than a 'typical' splicing reaction). Usually, RNA isolated from large amounts of extract does not form clean bands on a gel (right lanes, Figure 2A). A fortuitous discovery was that splicing extract made from another wild-type strain, derived from S288C (Brachmann et al., 1998), allowed for the extraction of very clean RNAs. (These strains appear to produce less 'lipo-protein scuz', which must be separated from the extract during preparation by an ultracentrifugation step). This strain was thus used for the rest of the RNA crosslinking experiments.

Splicing-dependent crosslinks

A timecourse of the splicing reaction before the crosslinking reaction revealed that many of the crosslinks were not dependent on splicing conditions (ATP and time at

25°C). However, two crosslinks do appear to be specific to splicing conditions. In Figure 3, the crosslink labelled (a) appears only after incubation in splicing conditions (+ATP) for 10 minutes and accumulates further with time. Crosslink (b) appears very faint in the absence of incubation under splicing conditions (0 minutes, +ATP), but accumulates with time, concurrently with the accumulation of lariat intermediate and lariat product. While crosslink (b) appears to accumulate in the absence of ATP, it is possible that this large band (in the wells) is actually a different crosslinked RNA species. Indeed, this large –ATP crosslink migrates differently from (b) on a 4% polyacrylamide gel (Figure 2). Although more experiments are needed to pinpoint their timing, the kinetics of formation of (a) and (b) crosslinks are consistent with an interaction forming after the first chemical step of splicing. It will be interesting to see whether they require the activity of Prp16, an RNA helicase that mediates a conformational change between the two steps (Schwer and Guthrie, 1992).

ATP, without incubation, also appears to affect some of the crosslinks (Figure 3). These could reflect interesting interactions with snRNAs or inter-molecular crosslinks. Since in most standard *in vitro* splicing reactions, only a fraction of the pre-mRNA substrate actually gets converted into splicing products, the functional relevance of these pre-mRNA products remains to be determined. One should test whether these crosslinks require splicing factors or intron consensus sequences. Because I was interested in spliceosomal interactions of the 3' splice site UAG during the second step of the splicing reaction, I focussed on the late-forming crosslinks (a) and (b).

Identifying the crosslinks

To obtain an initial identification of whether the crosslinks of interest were to U6 or U2 snRNAs, I used oligos complementary to snRNA sequences to specifically direct digestion by RNase H. If a crosslink is to U6 or U2 snRNA, its migration on the gel should change after this treatment. Under conditions that specifically digest U2 snRNA, the (a) crosslinked species disappears, suggesting that this crosslink may be to U2 (Figure 4A). Neither species is affected by conditions that digest U6 snRNA. Figure 4B shows a Nothern of the RNase H reactions to verify that U2 and U6 were indeed specifically digested.

DISCUSSION

This chapter describes the development of a powerful assay for probing interactions in the catalytic core. An intron-containing substrate for *in vitro* splicing reactions was generated with a specific modification at position –3 of the 3' splice site, using the directed RNA ligation technique of Moore and Sharp (Moore and Query, 2000; Moore and Sharp, 1992). The modification at position –3 is a photo-activatable crosslinking agent (4-thio-uridine) and a radioactive label for use in characterizing crosslinked species. The substrate forms two specific crosslinks under splicing conditions. The slow kinetics of formation suggests that these crosslinks may reflect interactions that form during late steps of splicing.

Previously observed protein crosslinks to the 3' UAG

In previous studies, several proteins have been observed to crosslink residues near the 3' splice site (Chiara et al., 1996; Chiara and Reed, 1995; McPheeters et al., 2000). Consistent with genetic observations and our model, Prp8 is one of these proteins. What is the function of the other proteins that crosslink? One of these proteins, Prp16, mediates an ATP-dependent conformational change at the 3' splice site (Schwer and Guthrie, 1992). The activity of Prp16 results in a stronger crosslink of Prp8 to residues near the 3' splice site (Umen and Guthrie, 1995). Another 3' splice site-crosslinking protein, Slu7, has been functionally implicated in distinguishing the correct 3' UAG from other 3'UAG sequences upstream or downstream (Brys and Schwer, 1996; Chua and Reed, 1999; Frank and Guthrie, 1992).

Consisting of only three nucleotides, the 3' UAG sequence is unlikely to define the 3' end of the intron by itself. Indeed, its distance from the branchpoint and the presence of a preceding poly-pyrimidine tract play important roles in recognition by the splicing machinery. Moreover, recognition probably involves multiple steps. In metazoans, the protein U2AF35 has been recently demonstrated to recognize the 3' UAG sequence early in spliceosome assembly (Merendino et al., 1999; Wu et al., 1999; Zorio and Blumenthal, 1999). This factor does not appear to be required for later steps and has no apparent yeast homologue. Conversely, factors associated with the U5 snRNP that crosslink the 3' UAG are required for the second step (reviewed in (Umen and Guthrie, 1995). Many of these crosslinks appear to require completion of the first step, and the activity of Prp16 (Umen and Guthrie, 1995).

Since recognition and choice of the 3' UAG must occur before catalysis, it is possible that these events are separable. Indeed, mutations in hypothesized RNA contacts to the 3' UAG do not have a dramatic influence (compared to distance from the branchpoint sequence) on 3'UAG selection in competition assays (Luukkonen and Séraphin, 1997). The extent to which RNA interactions that are important for catalysis contribute to 3' splice site selection remains to be determined.

It will be interesting to further pinpoint the timing of formation of the RNA crosslinks in this study, by blocking and unblocking the reaction at different steps (using mutant extracts and by immunodepletion followed by complementation of specific factors, such as Prp16). Towards this end, the protocol for synthesizing the photo-crosslinking substrate is described in detail in the Methods section, so that it can be easily resynthesized and used by others to further characterize both RNA and protein interactions of the 3' splice site.

Genetic interactions of the 3' UAG

Previously observed genetic interactions have identified spliceosomal RNA residues that may interact with the 3' UAG. These include residues in the essential U6 ACAGAG motif: A mutation in this motif, U6 G52**u**, can suppress mutations in the last residue of the 3' UAG (Lesser and Guthrie, 1993). A similar phenotype was observed for mutations in residue A25 of U2, which lies within the bulge between U2/U6 helix Ia and Ib (Madhani and Guthrie, 1994). Mutations in this bulge can also suppress specific

mutations in U6 G52 (Madhani and Guthrie, 1994). These results suggest that a tertiary interaction between U2/U6 helix I and the U6 ACAGAG motif plays a role in 3'UAG recognition.

A specific mutation in the last residue of the 3' UAG can be suppressed by a mutation in the first residue of the 5' splice site (Parker and Siliciano, 1993). Other mutations in the 3' UAG can be suppressed by mutations in Prp8 (Chapter 1). Lastly, a specific combination of mutations in the U6 ACAGAG motif with a 5' splice site and a 3' splice site mutation can restore splicing (Chapter 2). These results suggest that the 5' splice site and Prp8, in addition to the U6 ACAGAG and U2/U6 helix I, are also involved in 3' UAG recognition.

While the genetic results can attribute a function to RNA and protein residues in recognition, they cannot attribute a mechanism for recognition (such as a direct interaction). Thus the identification of physical contacts in the spliceosome is important complementary approach to identifying interactions that comprise the spliceosome's catalytic core. It will be highly informative to identify the RNA residues that participate in the two late-forming spliceosomal RNA crosslinks to position –3 of the 3' splice site UAG.

MATERIALS AND METHODS

I. Synthesis of crosslinking substrates

The strategy for synthesis of actin substrate containing photo-activatable uridine analogs at position –3 of the 3'SS UAG was originally developed by Professor David McPheeters at Case Western University (who kindly provided some constructs, described below), but was modified extensively. In hopes that someone may continue this work, the procedures I used are described in detail here.

RNA oligos

A 13-mer RNA oligo, 5'-ACACAACAAAAUA-3', was synthesized by Dharmacon, with the U either as the normal uridine, or 4-thio-uridine, or 5-iodo-uridine. The oligos are stored, with their 'ACE' 2' OH protecting groups in aliquats as frozen pellets in amber tubes at –20°C. After deprotection, an aliquat is stored in 5 mM DTT (which is important for maintaining the 4-thio-uridine reactivity).

In vitro transcribed RNA

The 5' and 3' ends of the actin substrate were synthesized through *in vitro* transcription reactions using T7 and SP6 RNA polymerase, respectively. The templates were generated by PCR, from the plasmid pACT8, using *Pfu* Polymerase, and gel purified. The oligos used to PCR the 5' T7 template (T7AC1 and SA1) and the 3' SP6 template (BSP1 and AC4) were gifts from David McPheeters.

Sequence of the 5' T7 product (348 nts):

5'GGGUCGACGGAUCCCCCUUUUAGAUUUUUCACGCUUACUGCUUUUUUCU
UCCCAAGAUCGAAAAUUUACUGAAUUAACAAUGGAUUCUGGUAUGUUCUA
GCGCUUGCACCAUCCCAUUUAACUGUAAGAAGAAUUGCACGGUCCCAAUU
GCUCGAGAGAUUUCUCUUUUACCUUUUUUUUACUAUUUUUCACUCUCCCAU
AACCUCCUAUAUUGACUGAUCUGUAAUAACCACGAUAUUAUUGGAAUAAA
UAGGGGCUUGAAAUUUGGAAAAAAAAAAAAAACUGAAAUAUUUUCGUGAU
AAGUGAUAGUGAUAUUCUUCUUUUAUUUGCUACUUGUCUCAUGUACUA-3'

Sequence of the 3' SP6 product (174nts):

5'GACAAAACAAUAACGGUUAUUGAUAACGGUUCUGGUAUGUGUAAAGC
CGGUUUUGCCGGUGACGACGCUCCUCGUGCUGUCUUCCCAUCUAUCGUC
GGUAGACCAAGACACCAAGGUAUCAUGGUCGGUAUGGGUCAAAAAGAC
UCCUACGUUGGUGAUGAAGGGGAAUUCCGC-3'

The RNA transcripts were gel-purified. The 3' SP6 transcript was treated with Calf Intestinal Alkaline Phosphatase, molecular biology grade, from Roche, in the presence of RNA Guard (inhibitor of RNase A) at 37°C for 2 hours, then extracted with phenol/chloroform/isoamylalcohol and ethanol precipated. (Yields after treatment were checked by agarose gel).

In a typical reaction, ~100 pmols of phosphatased 3' SP6 transcript was kinased with ~120 pmols gamma 32P-ATP (7000 Ci/mmol, crude, from ICN) (~800 microCuries--HOT) using T4 DNA kinase from NEB, incubated at 37°C for 2 hours. The reaction is then passed over a G-50 spin column, extracted with phenol/chloroform/isoamylalcohol, and ethanol precipitated.

3-way RNA ligations

The ligation procedure is based on the technique of Moore and Sharp, using T4 DNA ligase(Moore and Query, 2000). The bridging DNA oligo that hybridizes to the three RNAs to be ligated is BRCC.

BRCC:

5'CCGTTATTGTTTTGTCTATTTTGTTGTGTTAGTACATGAGACACAGTAG-3'

Annealing: (the ratios of RNAs and oligos are important for yields (Moore and Query, 2000)). A 24 microliter reaction containing 40 pmols kinased 3'SP6 RNA, 50 pmols 5' T7 RNA, 200 pmols RNA oligo, 20 pmols BRCC DNA bridging oligo, 40 mM NaCl, 10 mM Tris pH 7.4, 20 mM EDTA, and 5 mM DTT was incubated at 68°C for 3 minutes, then 25 C for 5 minutes. (DO NOT add RNA Guard when heating!)

Ligation: The annealed reaction was then diluted to 60 microliters with 1x T4 DNA ligase buffer (50 mM Tris-HCl (pH 7.5), 10 mM MgCl2, 10 mM DTT, 1-2 mM ATP, 25 μg/mL BSA), 20 units RNA Guard, and 4.5 units of T4 DNA ligase (from Ambion). The reaction was incubated from 4 to 6 hours at 30°C. Then extracted with phenol/chloroform/isoamylalcohol, and ethanol precipitated. The resuspended products were gel-purified. The 3' kinased SP6 RNA was ligated to the RNA oligo with 50 – 95% efficiency, but less than 5% was incorporated into the desired product for the 3-way ligation. (Heterogenous 3' ends for the 5' T7 RNA, a common artifact of *in vitro* transcription, is the expected culprit). Yet the final yield of ligated product after gel purification (~1.5 pmols) was enough for several crosslinking experiments.

II. Crosslinking

Splicing extract (from the strain BY4333 (Brachmann et al., 1998)) was made using standard procedures, and splicing exracts were performed using standard conditions. 30 to 50 microliter splicing reactions, containing 30-50 fmol of crosslinking substrate, (made by ligation, above), were incubated at 25°C in the presence or absence of 2 mM ATP (+2.5 mM $MgCl_2$), for 0 to 20 minutes. Samples were then placed on ice.

For crosslinking, samples were spotted in 9 microliter drops onto a pre-chilled (4 C) metal block (covered in parafilm). They were then irradiated with a UVP Hi-intensity UV lamp (model B-100AP), with a 365 nm filter, at a distance of ~ 8 cm, at 4 C for 10 minutes. (Pre-warm the lamp for 10 minutes first). This is a 100-Watt Mercury Vapor bulb. For 5-iodo-uridine, which is activated at a slightly lower wavelength (316 nm), I also tried crosslinking with a hand-held 312 nm UV-lamp (from Fisher Scientific), estimated power of 23 Watts. While no RNA-RNA crosslinks were seen for the 5-iodo-uridine modification, it may need a higher energy to activate efficient RNA crosslinking (D. Ryan and J. Abelson, personal communication).

The ability of the substrates in these assays to crosslink to proteins was also assessed. The crosslinked samples were treated with Rnase T1, then run onto a 7.5 % SDS-PAGE gel, and exposed to film. UV-specific and modification-specific protein crosslinks to both substrates were observed. (The 4-thio-uridine substrate was crosslinked with higher efficiency). Use of this substrate for mapping the crosslink on Prp8 is discussed further in the appendix.

RNA was purified from the crosslinking reactions by treatment with Proteinase K, then extracting three times with phenol/choroform/isoamylacohol, and ethanol prepitating. The resuspended samples were run on either 6% or 4% polyacrylamide RNA gels.

III. Oligo-directed RNase H mapping

RNA purified from the crosslinking reactions, above, was first by heating to 68 C and slow cooling to 37°C with 1 mM of one of the following anti-snRNA oligos, in the presence of 50 mM KCl, 10 mM Tris pH 7.4, and 10 mM EDTA annealed (in a 5 microliter volume). Then, 10 microliters of Rnase H mix was added (40 mM KCl, 40 mM Tris pH 7.4, and 10 mM MgCl2), and the reaction was incubated at 37°C for 2hours. Then extracted with phenol/choroform/isoamylacohol, and ethanol precipitated. The resuspended sample was split into 2 gels: One 6% or 4% polyacrylamide RNA gel to look for a shift in the mobility of crosslink bands, and the other for a Nothern to assess the extent and specificity of snRNA depletion.

Oligos:

U2 (76-91): 5' AAGGTAATGAGCCTCA 3' (oCC065)
U6 Block 1: 5' GGTTCATCCTTATGCTGGGGTTC 3'(from Stephen Rader)
U6 Block 2: 5' AAAC GGTTCATCCTTATGCTGGGGTTCTGCTGATGA 3' (from Stephen Rader)

U2 (30-45): 5' AACAGATACTACACTT 3' (oCC064), and
U6d1 5' ATCTCTGTATTGTTTCAAATTGACCAA 3'

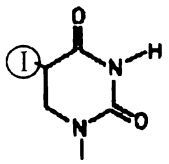
Figure 1: Strategy for identifying RNA interactions with the 3' splice site UAG.

A. Chemical structures, and the wavelength of light used for photo-activation, of the uridine analogs, 5-iodo-uridine, and 4-thio-uridine.

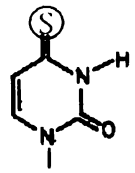
B. Outline of the strategy for substrate synthesis and crosslinking.

A

Photo-activatable
Uridine analogs:



5-iodo-uridine
316 nm



4-thio-uridine
366 nm

B

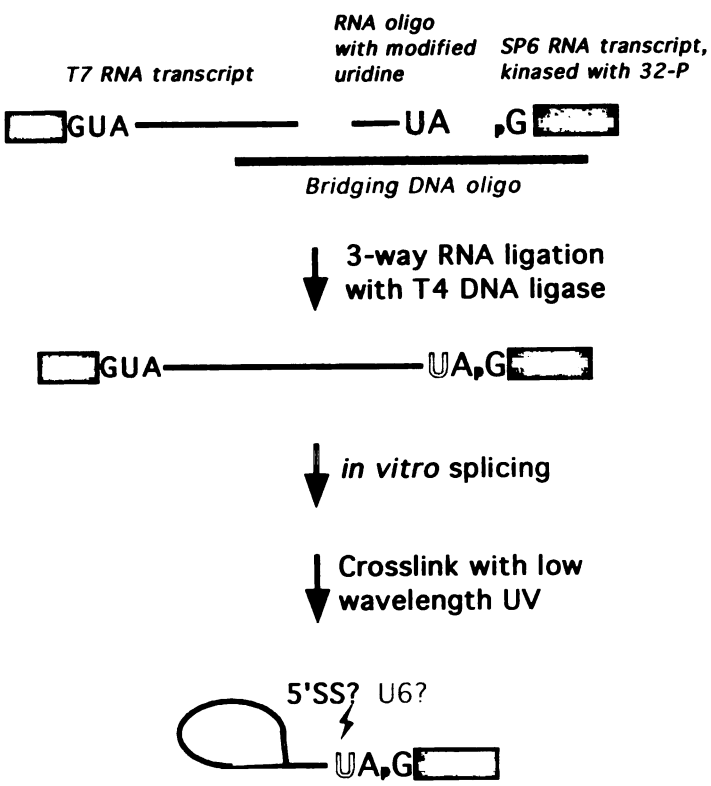


Figure 2: RNA crosslinks to position -3 of the 3' splice site.

A. S288C* denotes the 'designer' strain, described in (Brachmann et al., 1998), that allows for a much cleaner preparation of RNA and detection of RNA-RNA crosslinks. Splicing reactions with 5-I-U and 4-S-U substrates were incubated in the absence (-) or presence (+) of ATP for 20 minutes, then subjected to photocrosslinking (at 4°C) at 316 or 366 nm UV light (+), or kept on ice (-). RNA was extracted and run on a 6% denaturing polyacrylamide gel. The migration of precursor, lariat intermediate and excised lariat species is denoted on the side.

B. A 4% polyacrylamide gel allows for better separation of the RNA crosslinks. (a) and (b) crosslinks require incubation in the presence of ATP. (Also in Figure 3). The left lane shows a 100 base pair ladder.

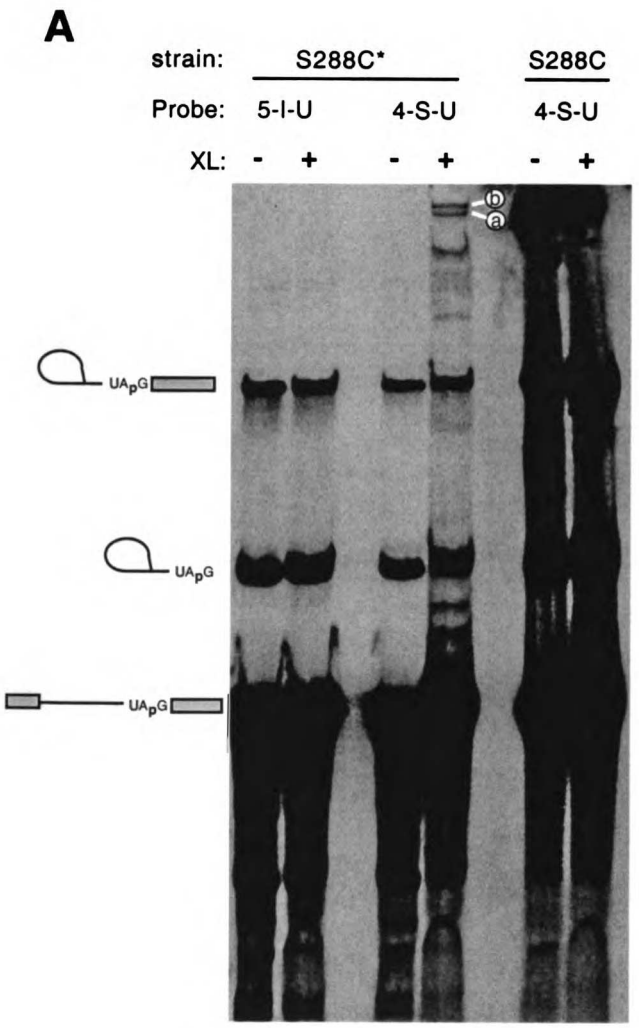

A
strain:
S288C*
S288C
Probe:
5-I-U
4-S-U
4-S-U
XL:
- + - + - +
b
a
UApG
UApG
UApG
6% polyacrylamide
B
ATP:
+
-
XL
:
- + - +
b
a
1,500
1,200
1,000
900
800
700
600
500
400
300
4% polyacrylamide

Figure 3: Time course of RNA crosslinks to position 3 of the 3' splice site.

Splicing reactions with the 4-S-U 3'splice site crosslinking substate were incubated for the indicated minutes at room temperature, in the presence (+) or absence (-) of ATP. RNA was then extracted and loaded on a 6% denaturing acrylamide gel (as in Figure 2A).

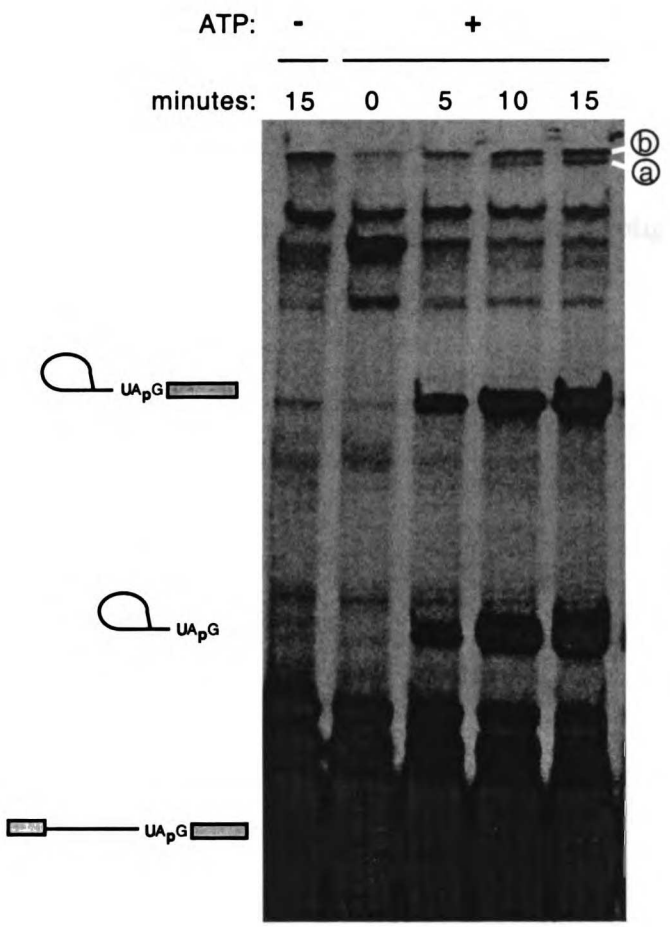
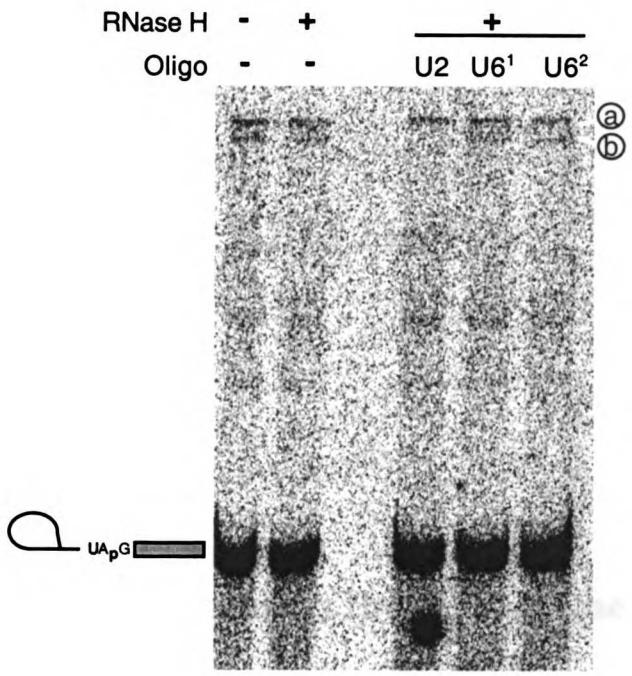

ATP: - +
minutes: 15 0 5 10 15
b
a
UApG
UApG
UApG
6% polyacrylamide

Figure 4: Identifying RNA crosslinks by oligo-directed RNase H cleavage.

A. Crosslinked species before and after treatment with oligos that hybridize to U2 or U6, and RNase H.

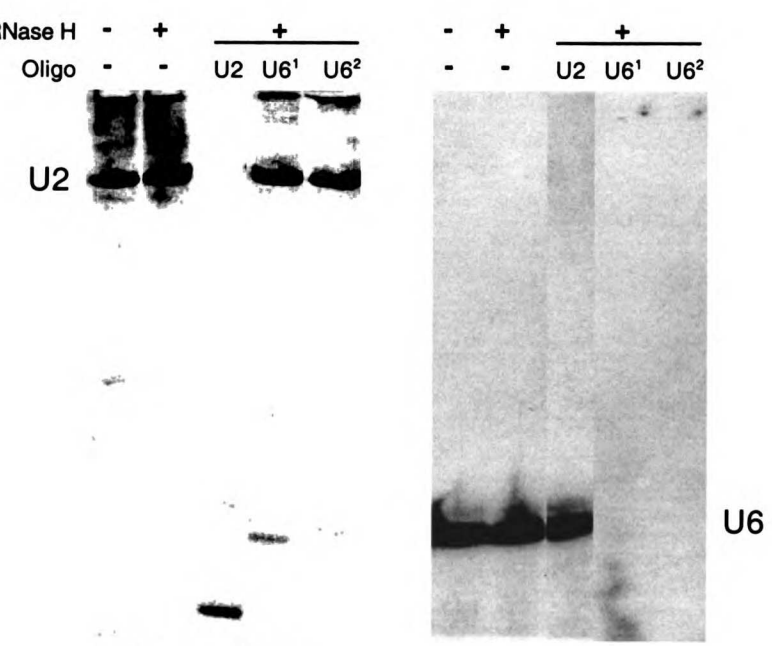
B. SnRNA Northerns to control for the extent and specificity of oligo directed cleavage.

A



6% polyacrylamide

B



U2 Northern U6 Northern

EPILOGUE

The question remains: Is the spliceosome a ribozyme?

The removal of introns by pre-messenger RNA (pre-mRNA) splicing is catalyzed by the spliceosome, a large ribonucleoprotein (RNP) machine comprised of at least 50 proteins and 5 small nuclear RNAs (U1, U2, U4/U6 and U5 snRNAs) (Burge et al., 1998). A longstanding question has been whether the spliceosome is fundamentally an RNA enzyme. As we will review, a network of essential RNA-RNA interactions could provide an RNA structure capable of catalyzing the two phosphoryl-transfer steps of the splicing reaction (Figure 1a).

But is there any role for protein in the chemical events of splicing? Although genetic analysis has demonstrated the requirement for dozens of spliceosomal proteins, it has been difficult to pinpoint specific mechanistic roles. Nonetheless, a single protein, Prp8, has emerged as a tantalizing candidate for a function at the heart of the spliceosome. We will review possible roles this protein may play in spliceosome catalysis.

Assigning individual roles to RNA and protein in catalysis will ultimately require structural analysis of the spliceosome's 'catalytic core'. Yet the dynamic nature of the spliceosome presents a daunting challenge. Unlike conventional enzymes, the spliceosome lacks a preformed active site. Instead, the catalytic core must be assembled anew upon each intron substrate, which entails multiple conformational rearrangements (Figure 1b). We will conclude with a discussion of recent progress towards a potentially powerful assay to identify the contributions of individual spliceosome components.

I. EVIDENCE THAT THE SPLICEOSOME IS AN RNA ENZYME

(1) Evidence that an RNA structure forms the catalytic core

Critical functions in catalysis include recognition and positioning of the substrate, and formation of an active site structure that can stabilize the transition state of the reaction. Work over the past ten years (reviewed in detail in (Madhani and Guthrie, 1994; Nilsen, 1998; Nilsen, 1994; Yu et al., 1998)) has identified a network of critical RNA-RNA interactions (Figure 2) that could possibly perform these functions. U6 and U2 interact with the intron near two of the three sites of chemistry, the 5' splice site and the branchpoint, via the *5'SS helix* and *BP helix*, respectively. The adjoining U2/U6 helices *I, II* and *III* could help juxtapose these two reactive sites for the first phosphoryl-transfer step. The highly conserved loop of U5 snRNA has been implicated in positioning the exons for ligation during the second phosphoryl-transfer step (Newman, 1997). Additional long-range interactions (dotted lines in Figure 2) have been identified that suggest the existence of a more compact RNA tertiary structure. These include an interaction between the first and last G's of the intron (a), and an invariant U6 residue near the *5'SS helix* with a bulged U2 residue in *helix I* (b).

This working model for RNA interactions in the catalytic core is derived from three kinds of analyses. (1) Phylogenetic comparisons of snRNA sequences from organisms as diverse as yeast, plants, mammals, worms, and trypanosmes (including major, minor and trans spliceosomal variants) reveal striking conservation in secondary structure. Notable are the RNA sequences which are absolutely conserved (shown in Figure 2); mutational analysis has confirmed that most of these residues are essential for

splicing. (2) Transient RNA-RNA interactions have been detected by various photo-crosslinking techniques. (3) Functional interactions between RNA residues have been demonstrated by compensatory genetic analysis, in which the splicing defect conferred by one RNA mutation is specifically rescued by a second RNA mutation. This approach is particularly informative when suppression is specific to the ability of the mutated residues to form a Watson-Crick basepair.

Interestingly, formation of the catalytic core appears to be regulated, and perhaps chaperoned, by the dynamic process of spliceosome assembly. Biochemical characterization of this assembly process has revealed a series of binding and destabilization steps between sub-spliceosomal particles (snRNPs) containing snRNA complexed with protein (Figure 1b). The structure of the RNA core (Figure 2, top) is dramatically different from the structures of the U2 and U6 snRNAs (Figure 2, bottom) in the snRNPs, and requires disruption of mutually exclusive interactions. The contributions of U6 to *helix I* and to the *intramolecular U6 helix* are excluded within the U4/U5/U6 triple-snRNP by extensive base-pairing interactions with U4. U2's contributions to U2/U6 *helix I* and *helix II* are excluded by an intramolecular stem in the U2 snRNP. A number of spliceosomal factors, including a family of ATPases (so-called DEAD-box proteins) that potentially catalyze RNA unwinding, have been implicated in mediating these dramatic rearrangements ((Staley and Guthrie, 1998) and references therein).

Additionally, there is at least one ATP-dependent conformational rearrangement between the two catalytic steps of splicing (Chua and Reed, 1999; Umen and Guthrie, 1995). Crosslinking data argue that tertiary interaction (c) (Figure 2) does not form until

the second step (Sontheimer and Steitz, 1993), and genetic data suggest that interactions (a) and (b) are not required until this point (Madhani and Guthrie, 1994; Parker and Siliciano, 1993). An important unanswered question is the degree of structural change in the catalytic core between the first and second chemical steps (Umen and Guthrie, 1995).

(2) The spliceosome is a metalloenzyme

For most RNA enzymes studied to date, (especially those that catalyze phophoryl-transfer reactions), metal ions are essential co-factors for catalysis(Narlikar and Herschlag, 1997; Pan et al., 1993; Yarus, 1993). It has recently been demonstrated that the spliceosome requires Mg^{2+} to bind the leaving group for catalysis for both phosphoryl-transfer steps (Gordon et al., 2000; Sontheimer et al., 1997).

Since metal ions are also essential co-factors for many protein enzymes, it is possible that protein side-chains in the spliceosome could function to bind and position these Mg^{2+} ions. However, several studies have implicated residues in U6 in binding Mg^{2+} (hatched circles in Figure 2) (Fabrizio and Abelson, 1992; Yu et al., 1995). This information comes from observations that substitution of phosphate oxygens at specific positions in RNA with phosphorothioate (which does not bind Mg^{2+}) interferes with the splicing reaction. A more convincing argument for a Mg^{2+} binding site can be made if addition of Mn^{2+} (a preferred ligand for sulfur) rescues the splicing defect of the phosphorothioate substitution. This has been recently demonstrated for one position within a region of highly conserved sequence in the U6 *intramolecular helix* (solid circle, asterisk in Figure 2) (Yean et al., 2000). Since metal ions can also function in stabilizing

RNA tertiary structure (Pan et al., 1993), an important remaining question is whether this bound Mg^{2+} plays a structural or catalytic role. Since the phosphorothioate substitution does not confer an obvious block to spliceosome assembly, a catalytic role for the RNA-bound Mg^{2+} is favored.

(3) Possible origin of the spliceosome from an RNA enzyme

A number of similarities have been noted between the spliceosome and Group II introns, a class of RNA enzymes that catalyze their own removal by a similar splicing reaction. The two phosphoryl-transfer steps of the splicing reaction proceed with the same stereochemistry (Moore and Sharp, 1993; Padgett et al., 1994), and the nucleophile for the first step is in both cases a 'branchpoint' 2' OH from a conserved adenosine within the intron. Other shared features (reviewed in more detail elsewhere, Michel and Ferat, 1995; Pyle, 1996; Weiner, 1993) include regions of secondary structures, sequences implicated in metal binding, and the function of a stem loop in positioning exons for ligation. Recently, a new tertiary interaction has been identified in Group II introns that juxtaposes an essential catalytic domain (D5) to the 5' splice site (Boudvillain et al., 2000). It will be interesting to see whether an analogous interaction takes place between the U6 intramolecular stem and the 5' splice site in the spliceosome. Regardless of the degree of evolutionary relationship between the spliceosome and any extant self-splicing RNA, the question remains how far diverged is the present-day machinery from an RNA-based catalytic strategy.

II. EVIDENCE THAT THE PROTEIN *PRP8* FUNCTIONS AT THE CATALYTIC CORE

If a spliceosomal protein were to play a critical role in the chemistry of splicing, evidence should be based on the same kinds of observations that implicate RNA at the catalytic core: crosslinking, phylogenetic conservation, and genetic interactions. While a number of proteins have been observed to crosslink the pre-mRNA substrate near the sites of chemistry (reviewed in Chiara et al., 1996), only Prp8 makes extensive crosslinking contacts near the sites of chemistry for both chemical steps. Prp8 also forms crosslinks within snRNPs to U5 (extensively) and U6 snRNAs (Dix et al., 1998; Vidal et al., 1999). The crosslinked residues are highlighted in Figure 2. Moreover, Prp8 exhibits striking phylogenetic conservation: 62% identity from yeast to man throughout its entire sequence of ~2000 amino acids. Unfortunately, Prp8's sequence is entirely novel, precluding testable predictions for its activities.

The third type of evidence would be the identification of allele-specific suppression between mutations in Prp8 and essential RNA residues. The position of an efficient crosslinking interaction with the /GU dinucleotide at the 5' splice site was mapped within human Prp8 (Reyes et al., 1999; Reyes et al., 1996). The corresponding region in the yeast protein was mutagenized, and alleles identified that could suppress the splicing defect due to mutations of the U at position 2 of the intron (Siatecka et al., 1999). In a parallel genetic screen in which the entire coding sequence of Prp8 was mutagenized, the region of the crosslink and three additional regions of Prp8 gave rise to

the same splice site suppression phenotype(Collins and Guthrie, 1999). The close proximity of the crosslinking site with a number of the suppressor mutations argues strongly that Prp8 functions in binding the 5' /GU.

Remarkably, these *prp8* alleles also suppressed intron mutations near the 3' splice site, and a mutation of the residue in U6 previously shown to cross-link to position 2 of the 5' splice site (Collins and Guthrie, 1999; Kim and Abelson, 1996; Siatecka et al., 1999; Sontheimer and Steitz, 1993; Umen and Guthrie, 1996). To explain these results, a model was proposed (Figure 3) in which Prp8 recognizes an RNA tertiary structure element comprising nucleotides near the 5' and 3' splice sites and U6. The model rationalizes other crosslinking interactions of Prp8 (to the 3' splice site and U6) and brings together residues that are highly conserved and essential. This provides a possible rationale for their requirement during the second catalytic step, and suggests that the proposed interaction involving Prp8 could be fundamental to the mechanism of splicing.

It thus appears that Prp8 acts *in conjunction* with RNA at the catalytic core. What might its activities with RNA be?

Evidence that Prp8 stabilizes tertiary RNA interactions

One plausible role for Prp8 is to stabilize the structure of the core. The tolerance of the splice site suppressor alleles for a number of mutations at specific RNA residues could arise from a loosening of constraints upon a tertiary RNA structure. A possible structural mechanism for this effect is provided by mutant versions of α-lytic protease and T4 lysozyme which display broadened substrate specificity. These mutant enzymes

have been found to exhibit an increased structural plasticity, based on crystallographic measurements of thermal motions in the substrate-binding pocket (Bone et al., 1989; Morton and Matthews, 1995).

Prp8 is also thought to stabilize interactions between variant exon sequences and the highly conserved loop of U5 snRNA, which are not likely to be strong enough alone to tether the cleaved 5' exon between the first and second chemical steps (Beggs et al., 1995; Newman, 1997). Interestingly, there are no other known RNA-RNA interactions that could juxtapose the U5 loop with the rest of the RNA core. Based on the extensive crosslinking interactions of Prp8 with U5, the 5' and 3' exon/intron junctions and U6, this protein is the most likely candidate to achieve this task.

Evidence that Prp8 facilitates formation of the catalytic core

In addition to providing stability, Prp8's interactions with RNA could conceivably promote formation of the catalytic core. Indeed, Prp8 interacts with several RNA components of the core prior to formation of the catalytically active spliceosome. Crosslinks to U6 (and U5) in the triple snRNP(Dix et al., 1998; Vidal et al., 1999), and to the 5'/GU at the stage of triple snRNP addition (Reyes et al., 1996)(Figure 2), imply that Prp8 is present during the RNA rearrangements required for catalytic activation. Through these early interactions, Prp8 might promote the folding of the active RNA structure.

A particularly interesting notion is that Prp8 may also facilitate formation of the core by coordinating the activity of spliceosomal factors that themselves catalyze RNA

rearrangements. Within snRNPs, Prp8 is tightly associated with two such factors (Achsel et al., 1998), a DEIH-box ATPase thought to unwind the extensive basepairing interactions between U6 and U4 (Raghunathan and Guthrie, 1998), and a GTPase homologous to the ribosomal translocation factor EF-2 (Fabrizio et al., 1997). The specific interactions of Prp8 with spliceosomal RNA could direct, and possibly regulate, the activity of these 'rearrangeases' (Kuhn and Brow, 2000).

Two lines of evidence suggest that Prp8 does *function* during catalytic activation. This step can be blocked genetically by: (1) a mutation in U4 (U4-cs1) that forms additional basepairs to U6 (shown in Figure 2)(Kuhn et al., 1999), which prevents interactions of U6 residues with the 5' with splice site; and (2) a mutation (*prp28*-1) in a DEAD-box ATPase, which inhibits the exchange of U1 for U6 at the 5' splice site (Staley and Guthrie, 1999). Strikingly, numerous mutations in Prp8 have been identified which can suppress the U4-cs1 mutation (Kuhn and Brow, 2000; Kuhn et al., 1999), and a distinct Prp8 mutation can suppress prp28-1 (Strauss and Guthrie, 1991).

III. DECONVOLUTING THE FUNCTIONS OF RNA AND PROTEIN IN SPLICEOSOME CATALYSIS

The preceding observations support an intimate functional interaction between Prp8 and RNA components of the catalytic core. While it remains possible that amino acids of Prp8 (or, formally, some other protein) make direct contributions to the catalytic mechanism (for example by coordinating a catalytic metal ion), the prevailing evidence (in section I) favors RNA as the catalytic component.

Thus a more likely possibility is that Prp8 acts as a protein co-factor to an RNA enzyme. Protein co-factors are required by many RNA enzymes, including most Group I and Group II introns, and RNase P, for catalysis under physiological conditions. However protein-independent catalysis can occur in conditions that stabilize RNA structure, such as a high concentration of divalent metal ions, or, as in (Weeks and Cech, 1995), the presence of a crosslink that stabilizes a long-range tertiary interaction. Well-studied examples have generally assigned functions to a protein co-factor in stabilizing the catalytic RNA structure, and in facilitating its formation (Weeks, 1997). Evidence for such functions by Prp8 have been described above.

As a third possibility, RNA and protein might function intimately together in catalysis through a specific protein-RNA structure that cannot be acheived by RNA alone. Many proteins that bind specifically to RNA have been observed to induce dramatic conformational changes in the RNA, or to stabilize a unique RNA configuration (Frankel and Smith, 1998; Weeks, 1997). Interesting recent examples include the dramatic distortion of SRP RNA by its highly conserved protein co-factor (Bernstein,

2000); and the specific effects of the ribosomal protein L11 upon the stability a 23S rRNA tertiary structure through direct amino acid contacts (Conn et al., 1999).

How might the possible contributions of Prp8 and RNA to folding, structural stabilization and catalysis be distinguished? Unfortunately, the detailed enzymatic and structural studies that have elucidated protein function in the reactions of RNA enzymes are not directly applicable to the spliceosome. Because the catalytic steps are not easily uncoupled from the multiple assembly steps, kinetic parameters (such as rate constants and binding affinities) for individual steps cannot be easily measured. In the ideal case, one would like to engineer a system in with the catalytic core could be studied in isolation.

Towards reconstitution of a catalytic core

Recent studies have allowed promising advances towards such a reductionist approach. Synthetic transcripts corresponding to the regions of U2 and U6 shown in Figure 2 (top) were annealed by heating and slow cooling in the presence of high concentrations of Mg2+ (Valadkhan and Manley, 2000). Remarkably, the U2-U6 tertiary interaction identified genetically in the spliceosome (b), can be detected by crosslinking in this annealed complex. Experiments in progress have verified the existence of the predicted secondary structure elements. The U2-U6 complex also displays an ability to bind sequences derived from the 5' splice site and branchpoint consensus (S. Valadkhan and J. L. Manley, personal communcation). The ultimate question is whether sufficient

interactions can be built upon this complex to form a structure that can accurately catalyze one of the spliceosome's phosphoryl-transfer reactions .

In vitro assembly of a catalytic core could be a powerful assay for dissecting the functional and structural contributions of its components, including Prp8. With such a system, Prp8 might be assayed for effects on the formation and the stability of an RNA tertiary structure, and possibly upon the function of this structure in catalysis.

The example of the ribosome

In conclusion, as the field faces the enormous challenges presented by the dynamic nature of the spliceosome, we can look to the ribosome, the cell's other large RNP, with hope and awe. Exciting advances in ribosome crystallography have recently revealed an atomic resolution structure of the fully formed peptidyl-transferase active site, crowning more than three decades of study, and verifying RNA as the catalytic component(Ban et al., 2000; Nissen et al., 2000). The time seems ripe to tackle the spliceosome's catalytic core.

ACKNOWLEDGEMENTS

We thank John Abelson, Amy Kistler, Maki Inada, Jon Staley, Stephen Rader and Richard Wagner for helpful comments and discussions.

This chapter was published in *Nature Structural Biology*, vol. 7: 850-4.

Figure 1: Spliceosome Reaction and Assembly.

A. The pre-mRNA splicing reaction consists of two phosphoryl-transfer steps. In the first step, the 5' intron phosphate, (the 5' splice site) is attacked by a 2' hydroxyl specified within the intron (the branch point). In the second step, the 3' intron phosphate (the 3' splice site) is attacked by the 3' hydroxyl of the cleaved 5' exon. The final products of the splicing reaction are ligated exons, and the excised intron in a branched, 'lariat' form.

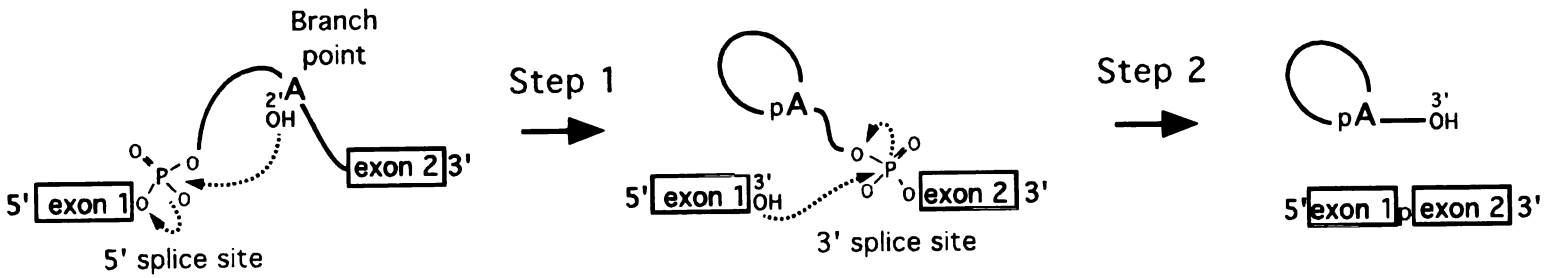

Branch
point
2'
OH
A
5'
exon 1
exon 2
3'
5' splice site
Step 1
pA
3'
OH
3' splice site
Step 2

Figure 1: Spliceosome Reaction and Assembly.

B. Stages of spliceosome assembly.

Small nuclear ribonuclear protein particles, SnRNPs, (U1, U2, U4, U5, U6) assemble onto the intron substrate in a step-wise fashion, directed by interactions with intron consensus sequences. (The degree of conservation in sequence is denoted by size. Y denotes pyrimidine). The U1 and U4 snRNPs become destabilized during catalytic activation. Thus the catalytically active spliceosome, consists of U2, U5, and U6 snRNPs.

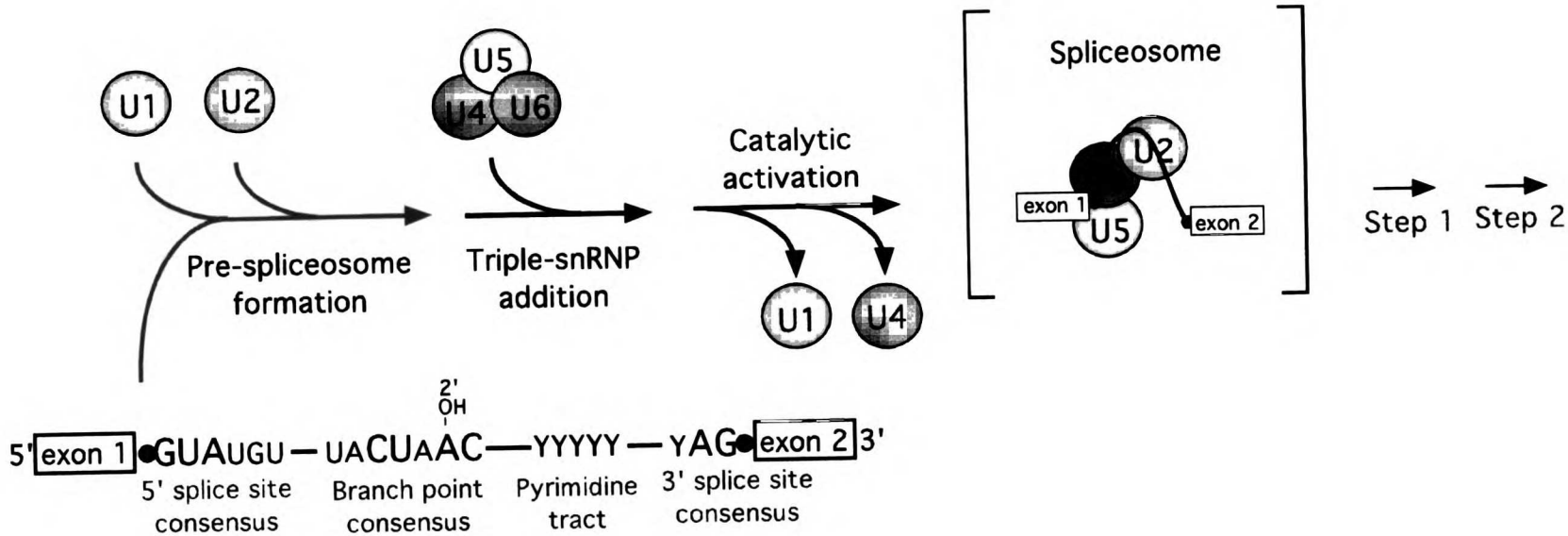

U1
U2
U5
U4
U6
Pre-spliceosome formation
Triple-snRNP addition
Catalytic activation
U1
U4
Spliceosome
U2
exon 1
U5
exon 2
Step 1
Step 2
2'
OH
5' exon 1 GUAUGU — UACUAAC — YYYYY — YAG exon 2 3'
5' splice site consensus
Branch point consensus
Pyrimidine tract
3' splice site consensus

Figure 2: Working model of RNA and Prp8 interactions in the catalytic core, and mutually exclusive interactions of U6 and U2 snRNAs in pre-assembled snRNPs.

Large letters denote RNA sequences that are absolutely conserved in major, minor and trans spliceosomes from mammals, worms, plants, yeast and trypanosomes. The exception, (underlined), terminal intron Gs are replaced by covariant 5' C and 3' A in some introns(Sharp and Burge, 1997). Black lines denote Watson-Crick basepairing interactions. (The thinner lines denote interactions that are not absolutely conserved in all systems). Exons are denoted by rectangles, and the intron substrate is in black. Purple dotted lines indicate tertiary interactions (a), (b), and (c). The smaller hatched circles indicate residues for which phosphorothioate substitution conferred a partial block to splicing; for the larger circles, the block was more severe (Fabrizio and Abelson, 1992; Yu et al., 1995). The phosphorothioate substitution that is recued by Mn^{2+} is indicated by the solid circle and asterisk.

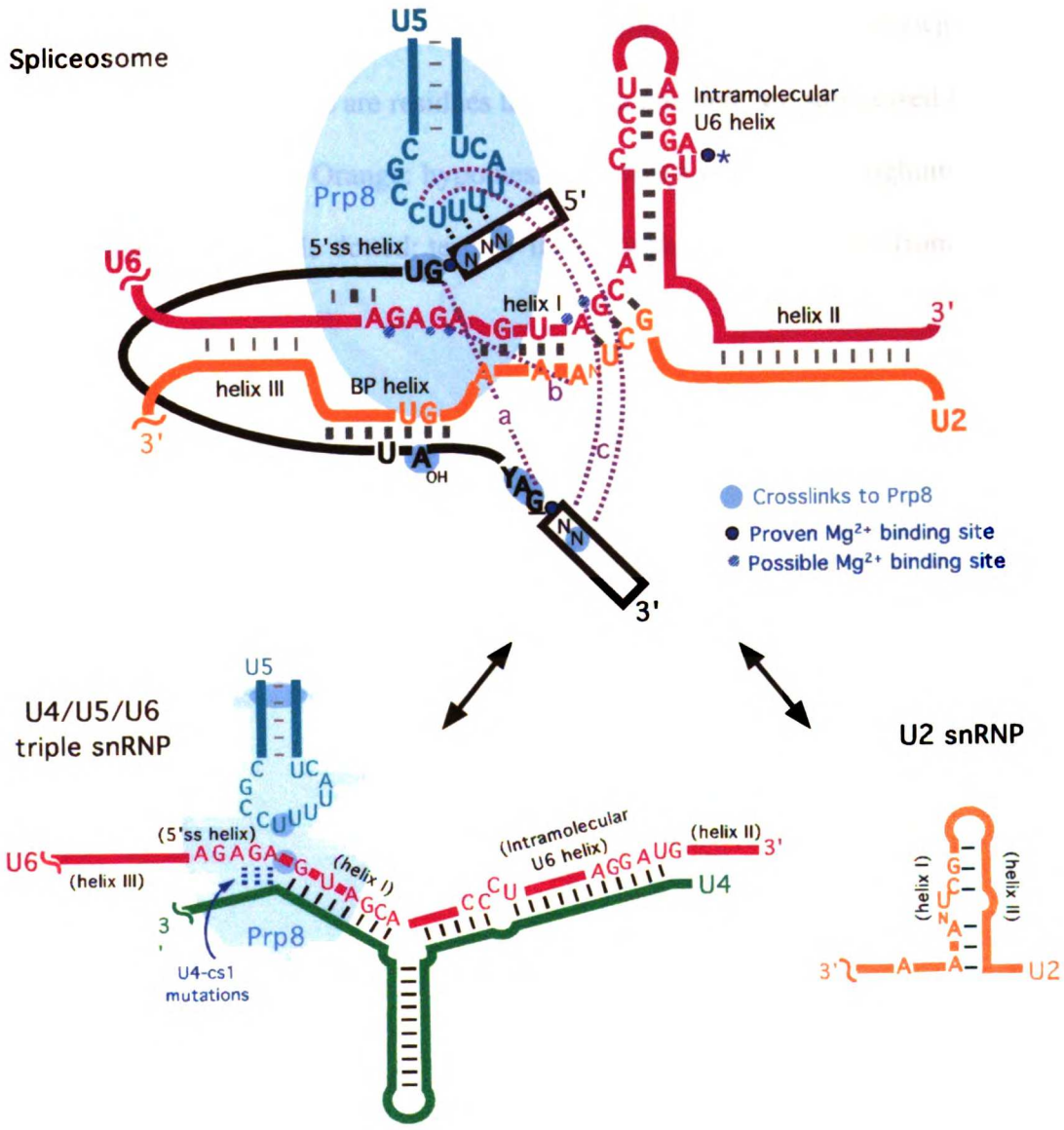
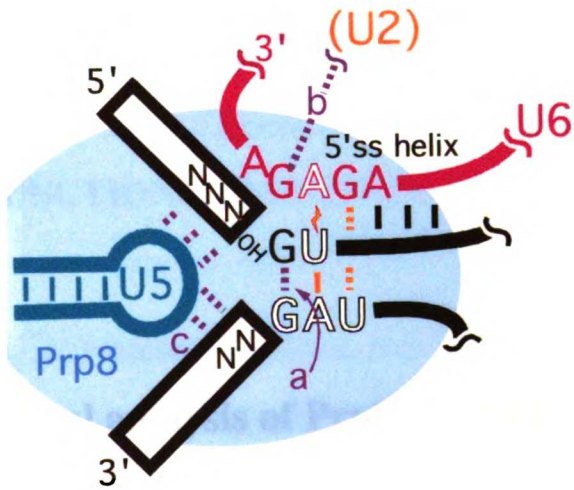

Spliceosome
U5
Prp8
5'
5'ss helix
U6
helix I
helix II
3'
U2
Intramolecular
U6 helix
helix III
BP helix
3'
OH
Crosslinks to Prp8
Proven Mg2+ binding site
Possible Mg2+ binding site
U4/U5/U6
triple snRNP
U5
(5'ss helix)
U6
(helix III)
(helix I)
Prp8
3'
U4-cs1
mutations
(Intramolecular
U6 helix)
(helix II)
U4
U2 snRNP
(helix I)
(helix II)
U2

Figure 3: Model: Prp8 binds a tertiary RNA structure in the catalytic core.

Some of the interactions of the active spliceosome described above (the 5'SS helix of U6 with the intron, and interactions of the U5 conserved loop with exons) are drawn for the second chemical step. In outline are residues that, when mutated, are suppressed by prp8 splice site suppressor alleles. Orange: hypothesized tertiary interactions; lightning bolt: crosslinking interaction; purple dotted: tertiary interactions (a), (b), and (c) from Figure 2.


5'
3'
(U2)
b
U6
5'ss helix
AGAGA
NNN
OH
GU
U5
GAU
Prp8
c
N
N
a
3'

APPENDIX

Towards biochemical analysis of Prp8's interactions with RNA

We have established that Prp8 is a fundamental protein component of the catalytic core. Discussed in the Epilogue, its function may be to chaperone or stabilize an RNA structure, or perhaps directly contribute to catalysis. Resolving these possibilities, and understanding the catalytic mechanism of pre-mRNA splicing, requires detailed structural analysis of Prp8 and its interactions with RNA at the catalytic core.

FUNCTIONAL ASSAYS FOR PRP8

Prp8 has yet to be recombinantly expressed, or purified and assayed for any biochemical activity. Assessing whether purified or recombinant Prp8 and Prp8 fragments are functionally active requires a splicing complementation assay. Because endogenous Prp8 is tightly bound in snRNPs, this would require conditions that allow for disruption, followed by reformation, of the snRNP interactions. Since the U5 snRNP alone consists of at least 20 proteins, this is no small task. A possible approach is to treat splicing extract with a high concentration of salt in order to disrupt the snRNP interactions, followed by removal of salt by dialysis to allow for snRNP reconstitution. The splicing activity of a whole cell yeast extract can survive a treatment with 500 mM NaCl (Staley and Guthrie, 1999). Because the association of Prp8 with U4 and U6 snRNAs is resistant to at least 550 mM NaCl (unpublished results), other conditions for disrupting the snRNP interactions would need to be tested.

A second strategy, which may more directly uncover Prp8's contibutions to the catalytic core, is to reconstitute Prp8's spliceosomal interactions from the ground up,

starting with synthetic RNAs. Discussed in the Epilogue, recent data provides great hope that the catalytic core, or at least parts of it, may be assembled *in vitro* (Valadkhan and Manley, 2000). The U2/U6 intramolecular helices, and the U6 intermolecular helix can be annealed in vitro by heating and slow cooling synthetic RNAs in the presence of Mg^{2+}. While it is not surprising that base pairing interactions form, crosslinking assays detect a tertiary interaction in this minimal system that resembles an interaction proposed in the spliceosome from *in vivo* covariation analysis (Madhani and Guthrie, 1994). The next step is see whether other known interactions of the catalytic core can be built upon this structure. For instance, can this complex bind a 5' splice site or branchpoint consensus sequence? With such a substrate, could catalysis of a phosphoryl-transfer reaction be observed? Does Prp8, or a fragment thereof, influence the efficiency of any of these activities?

The Manley lab has recently communicated that an annealed U2/U6 structure can indeed bind to a branchpoint consensus sequence, and can promote a phosphoryl transfer reaction using the branchpoint adenosine (!). Thus *in vitro* assembly of a catalytic core may be feasible assay for dissecting the functions, structure, and interactions of the RNA components and of Prp8. Tommaso Villa, a post-doc in the lab, is currently pursuing several approaches towards this compelling goal.

HOW DOES PRP8 INTERACT WITH RNA?

The above experiments suggest some possible functional assays for Prp8, which could be used in structure-function analysis of *how* Prp8 binds RNA. No RNA-binding motif or domain has yet been identified. As a large protein, Prp8 probably has a number of functionalities and interactions, which might be carried out by separable domains. It is also possible that some of Prp8's interactions require the cooperative activity of multiple domains.

Summary of genetically identified functional domains

Thus far, multiple yet discrete regions of the protein have been genetically implicated in distinct functions. Mutations that give rise to distinct Prp8 phenotypes, for the most part, map to non-overlapping regions of the protein (Figure 1). The locations of *prp8* splice site suppressor alleles (which suppress mutations in specific residues near the 5' splice site, 3' splice site, and in U6 snRNA) are indicated in purple. Light blue shading indicates the locations of *prp8* mutations that can suppress the cold-sensitive growth defect due to a mutation in U4 snRNA (U4-cs1) (Kuhn and Brow, 2000; Kuhn et al., 1999). By forming extra base pairing interactions with U6 snRNA, this U4 mutant appears to prevent U6 from pairing with the 5' splice site. The mechanism of suppression by Prp8 is not known, but it is intriguing to speculate that this has something to do with Prp8's interaction with the ATPase Brr2, which has been implicated in U4/U6 unwinding (Raghunathan and Guthrie, 1998), or with Prp8's interactions with the 5'

splice site and U6. Interestingly, one allele, (*prp8-201*) displays both splice site suppression and U4-cs1 suppression phenotypes (data not shown). Perhaps the distinct phenotypes indeed share some mechanistic basis. It will be interesting to further characterize these alleles with respect to binding assays with RNAs and other snRNP proteins.

Green shading indicates the location of *prp8* mutations that confer an inability to utilize the poly-pyrimidine tract for selecting the correct 3' splice site. The mechanism for this phenotype is also not known. Intriguingly, the poly-pyrimidine *prp8* mutations map to the same general region of the protein as some of the splice site suppressor and U4-cs1 suppressor alleles.

In summary, the genetic analysis suggests that multiple regions of Prp8 can contribute to one function. Some regions of the protein are implicated in distinct functions, and a few regions of the protein are implicated in multiple functions. Hopefully, this functional map will complement future biochemical and structural characterizations of the domain organization and interactions of Prp8.

Future experiments to biochemically map Prp8 domains

Alan Kutach, a post-doc in the lab, is taking a number of approaches towards identifying independently-folding, functional units of Prp8. These include screens of Prp8 libraries for fragments that confer a dominant negative phenotype, or that can complement each other in *trans*, and using tranposons to generate libraries of *prp8* mutants that have substitutions scanning the protein.

INITIAL INROADS

Mapping Prp8's RNA contacts

We are most interested in identifying the domain or domains of Prp8 that physically contact RNA. Towards this end, the location of the crosslink to the 5' splice site GU has been mapped to within 10 amino acids on Prp8, using a protease fingerprinting approach (Reyes et al., 1999). Denoted by the lightening bolt in Figure 1, the crosslink falls within a region delineated by the location of splice site suppressor alleles, consistent with the model that Prp8 influences an RNA structure via a direct RNA-protein intearction. The crosslink also lies close to residues implicated in other phenotypes. It will be interesting to see whether other RNA contacts, for instance, to the 3' splice site, U6 snRNA, and U5 snRNA, map to similar or different regions. The model that Prp8 binds a tertiary RNA structure predicts that different RNA crosslinks will map to the same region of Prp8. However, the genetic results suggest that RNA-binding might be carried out by a cooperation between multiple domains. RNA crosslinks at different locations in the primary sequence of Prp8 could still reflect interactions that are close in 3-dimensional space.

A general strategy for mapping RNA crosslinks on Prp8 is outlined in Figure 2. This takes advantage of epitope tags that can be generated on either the N or C-terminus of Prp8, and techniques for generating fragments of the tagged protein by partial digestion with various proteases. By using an N or C-terminal epitope tag to purify

crosslinked fragments of Prp8, and identifying the minimal size of the crosslinked N-terminally or C-terminally tagged fragment, one can roughly estimate the boundaries for where the crosslink must be. High resolution mapping of the actual amino acid contacts would require much more time and material. The relevant information to be obtained by this approach is simply whether different RNA crosslinks involve the same or distinct folding domains.

Described in Chapter 3, the 3' splice site crosslinking substrate for probing contacts of position –3 (the U of the UAG), is predicted to crosslink to Prp8. Indeed, a highly efficient crosslink of the approximate size of Prp8 is detected under splicing conditions (not shown). The mapping procedure outlined above could thus be applied to this crosslink.

Identifying protease-resistant fragments of Prp8

When treated with proteases under non-denaturing conditions, the 3-dimensional structure of a protein typically protects some regions of the protein from cleavage by the protease, and exposes others. Identifying the sensitive and resistant regions can thus give a crude idea of the domain organization.

Preliminary results from a time-course of digestion with trypsin reveals that a ~40 kDa C-terminal fragment of Prp8 is protected (Figure 3). This is a standard size for an independently folding protein domain. A plausible site of cleavage could be the arginine at position 2105 which lies within a stretch of 4 residues (in yeast Prp8) that are not conserved between the known Prp8 homologues. There are 25 theoretical trypsin

cleavage sites (lysines and arginines) C-terminal to 2105 that appear to be protected from cleavage.

This C-terminal fragment shows 45 % identity between yeast to human Prp8, in contrast to 61 % identity for the entire sequences of the two homologues. A fragment of approximately the same size interacts in a 2-hybrid assay with Brr2 (J. Beggs, unpublished communication). The 2-hybrid interaction is supported by the observation that Prp8 and Brr2 form tight (salt-resistant) associations within purified U5 and U4/U5/U6 snRNPs (Achsel et al., 1998). This fragment also contains residues implicated in the splice site suppression phenotype. Each of the three mutations that map here confer a weak 5' and 3' splice site suppression phenotype, which, interestingly, appears stronger for the first step than for the second step of splicing (Chapter 1 and data not shown). It will be interesting express this fragment by itself and test whether it can form spliceosomal interactions in *trans*.

A Prp8-centric assay for conformational changes during splicing

As a component of the U4/U5/U6 triple snRNP, and of the spliceosome throughout assembly and both steps of catalysis, Prp8 is likely to be privy to many conformational rearrangements. Since many of these rearrangements involve RNA residues that comprise the catalytic core, (some of which interact with Prp8), Prp8 may be an important focal point. Formation or disruption of interactions, either within Prp8, or between Prp8 and other factors, could change the pattern of protease resistant and sensitive sites on Prp8.

Use of N-terminal or internal tags (which are being generated by Alan Kutach), and additional proteases, should allow us to identify additional resistant and sensitive sites in the protein. Such a map of Prp8 might not only provide hints of candidate domains, but might also provide a tool for detecting conformational changes during the course of splicing. If indeed regions of Prp8 change in protease sensitivity through the course of splicing, the protease map can be developed into an assay to characterize and compare mutations, both in Prp8 or in other splicing factors, that block splicing at various steps.

Figure 1: Current map of functionalities on Prp8.

The locations of mutations that give rise to the splice site suppression phenotype (purple) (Chapter 1, and (Siatecka et al., 1999; Umen and Guthrie, 1996)), U4-cs1 suppression phenotype (blue) (Kuhn and Brow, 2000; Kuhn et al., 1999), and poly-pyrimidine tract recognition phenotype (green) (Umen and Guthrie, 1996; Umen and Guthrie, 1995) have been mapped to the shaded regions. Dotted lines connect locations of U4-cs1 mutations that rescue or dramatically exacerbate each other when combined (Kuhn and Brow, 2000). Orange lines indicate direct and indirect 2-hybrid interactions with components of the U1 snRNP (Abovich and Rosbash, 1997; Kuhn and Brow, 2000) and S. Ruby, unpublished communication. The pink line indicates a fragment of Prp8 that interacts with Brr2 in a a two-hybrid assay (J. Beggs, unpublished communiciation). This two-hybrid interaction is inhibited the dotted pink region of Prp8 is included (J. Beggs, unpublished), suggesting that Prp8's interactions may be regulated by conformational changes within the protein. The C-terminal fragment that is resistant to trypsin digestion (Figure 3) is boxed in blue. The location of Prp8 of the crosslink to the 5' splice site GU dinucleotide has been mapped to region of 10 amino acids indicated by the lightening bolt (Reyes et al., 1999).

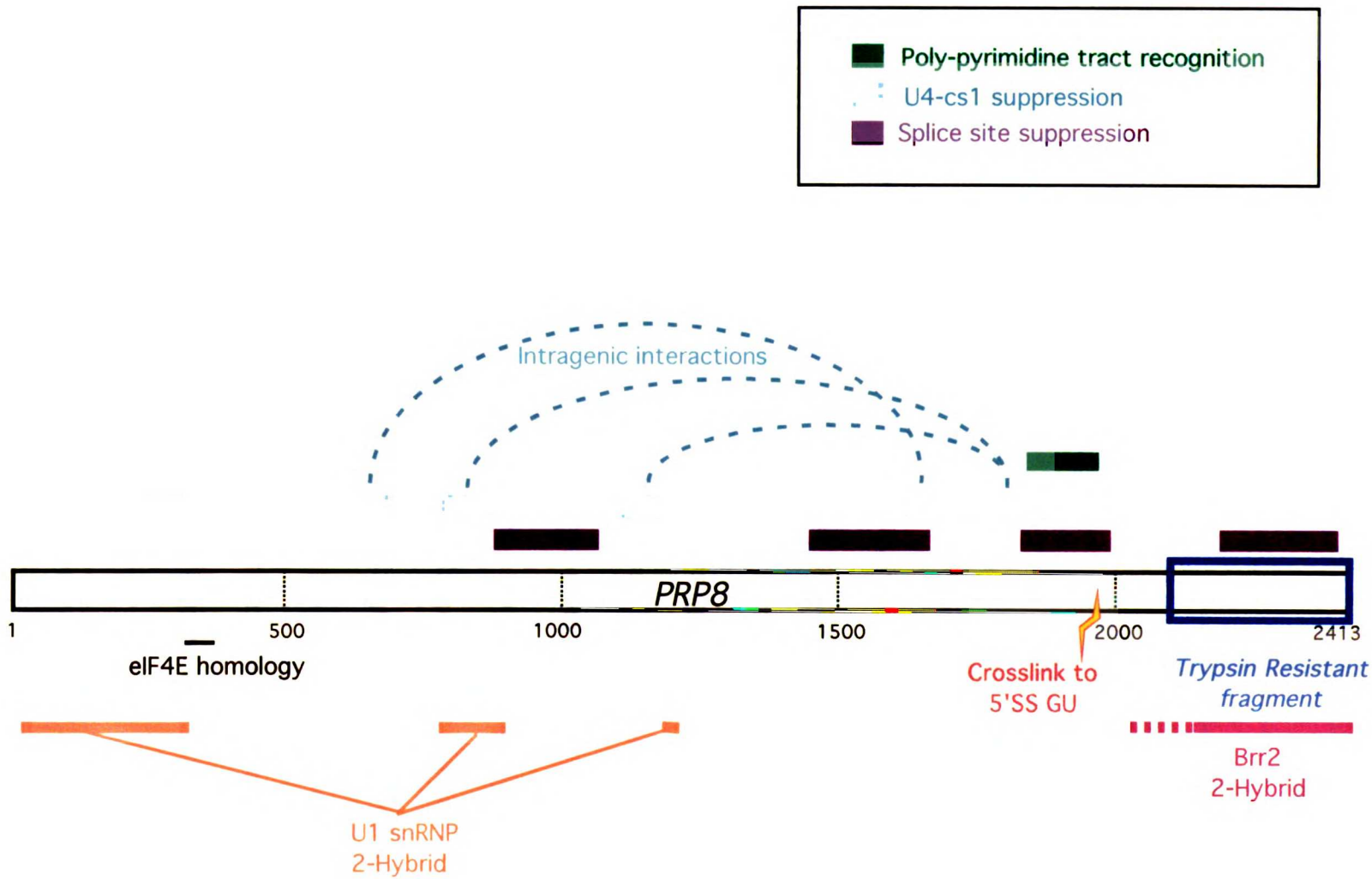

Poly-pyrimidine tract recognition
U4-cs1 suppression
Splice site suppression
Intragenic interactions
PRP8
1
500
1000
1500
2000
2413
eIF4E homology
Crosslink to
5'SS GU
Trypsin Resistant
fragment
Brr2
2-Hybrid
U1 snRNP
2-Hybrid

Figure 2: Strategy for mapping crosslink to fragment of Prp8.

First, an RNA-crosslinking reaction is performed upon splicing reaction that contains a specifically labeled RNA (see chapter 3). The extract is then treated with various proteases, such as trypsin, under conditions that allow for partial proteolysis (as in Figure 3). N-terminal and/or C-terminal epitope tags on Prp8 will allow for purification and detection of the partially digested Prp8 fragments. Determining which fragments contain the radioactive label from the crosslinked RNA should identify N-terminal and C-terminal boundaries for where the crosslink could lie on the protein. It is important for this strategy to purify N-terminally and C-terminally tagged fragments away from other Prp8 fragments. To do this, (indicated by asterisks, and described in Umen and Guthrie, 1995), one can completely denature Prp8 with SDS, then titrate the SDS by dilution with Triton X-100, in order to perform the immunoprecipitation. (If tags of repeated Histidines are used, this denaturing-IP step could be skipped, because the poly-His tag should bind tightly to a nickel column under denaturing conditions).

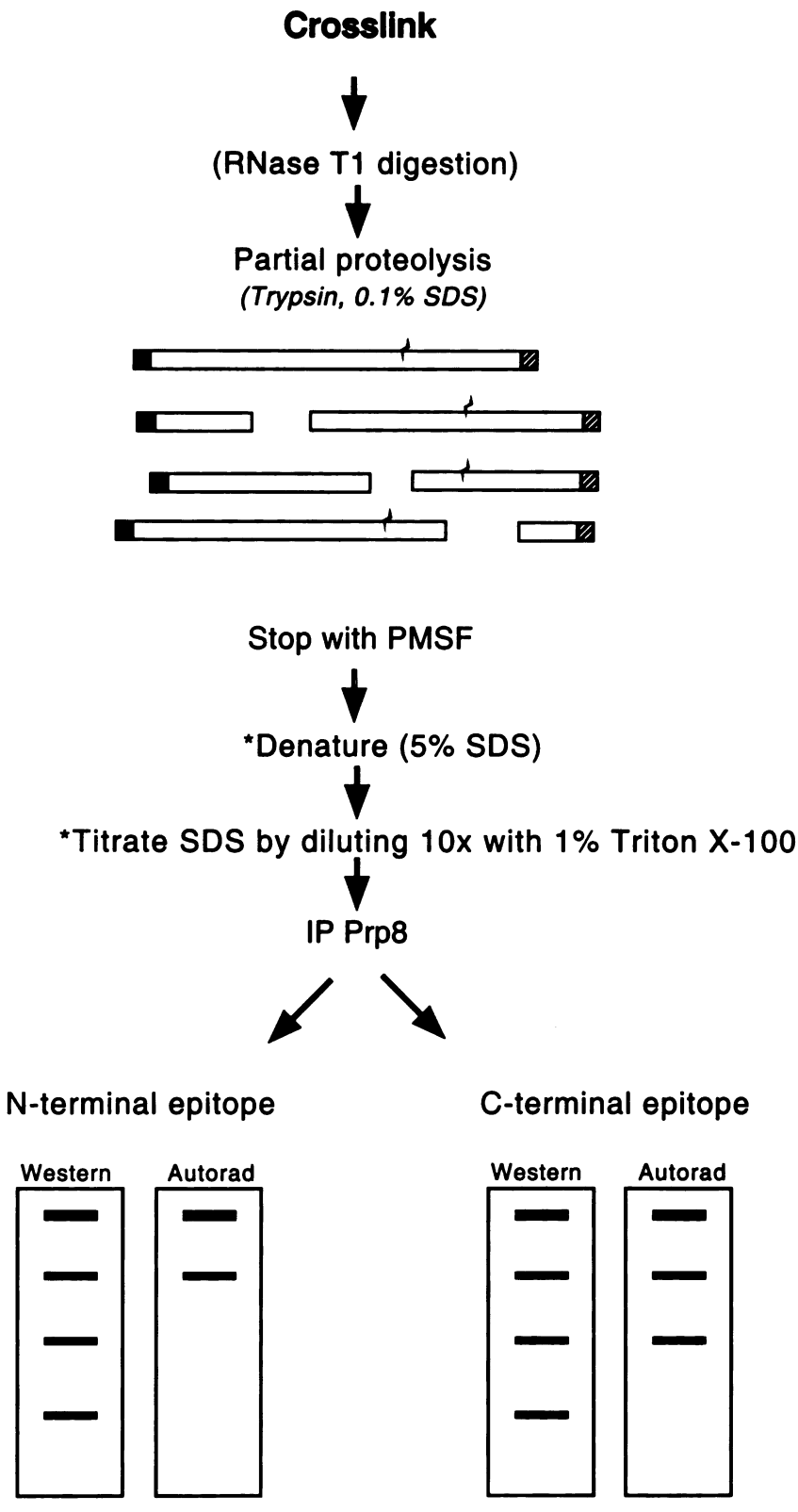
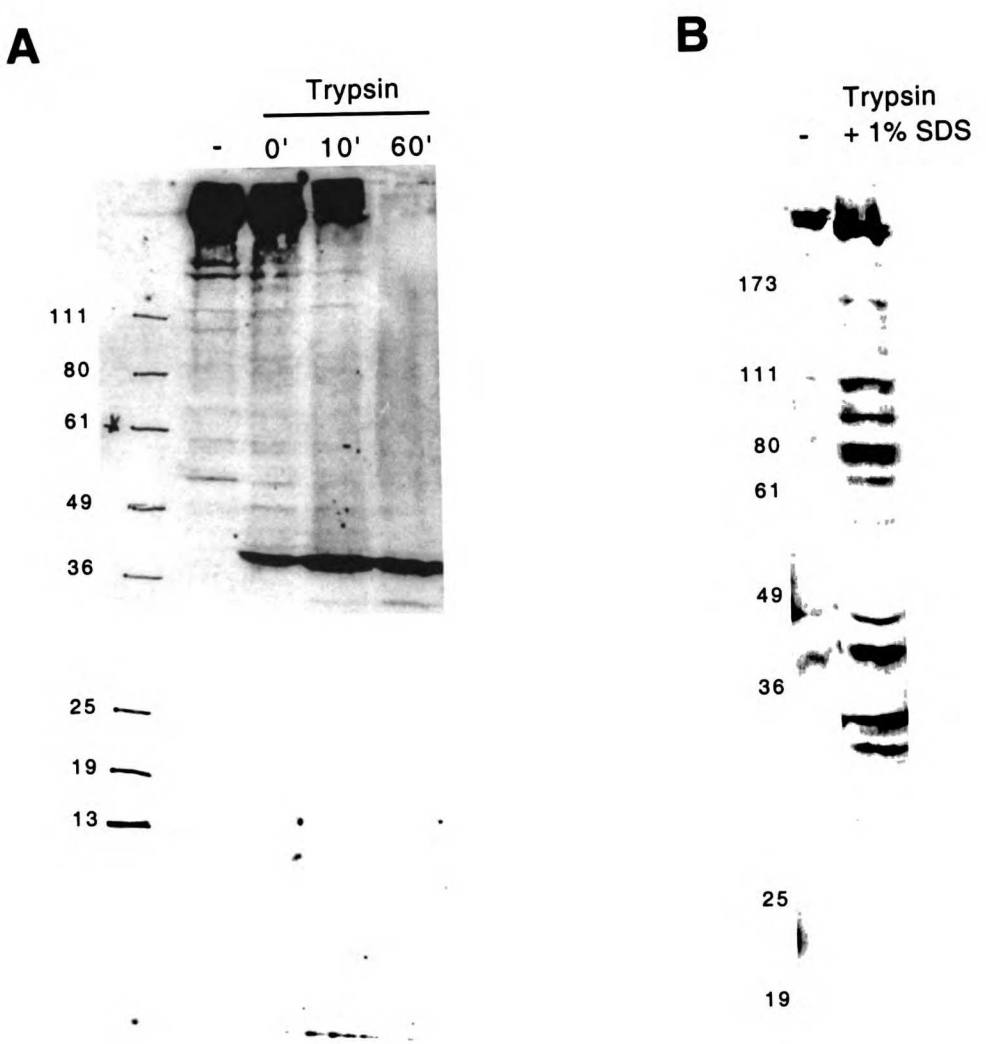

Crosslink
(RNase T1 digestion)
Partial proteolysis
(Trypsin, 0.1% SDS)
Stop with PMSF
*Denature (5% SDS)
*Titrate SDS by diluting 10x with 1% Triton X-100
IP Prp8
N-terminal epitope
C-terminal epitope
Western
Autorad
Western
Autorad

Figure 3: Trypsin digestion of Prp8(HA)$_3$

A. A C-terminal fragment of Prp8 is protected from trypsin in yeast splicing extract. 100 μg of splicing extract was diluted into 10 μL reaction containing 50 mM Tris-HCl pH 8.5, and, where indicated, 1 μg trypsin (modified, sequencing grade, from Roche), and incubated at 37°C for 0, 10, or 60 minutes before 'quenching' with 10 mM PMSF and SDS-PAGE sample buffer. Samples were run on a 4-20% gradient SDS-polyacrylamide gel, and blotted with 12ca5 monoclonal anti-HA antibodies. The migration of pre-stained molecular weight marker proteins is indicated on the left.

B. A partially digested ladder of Prp8(HA)$_3$ fragments is obtained by digestion with trypsin, using the conditions in part A for 60 minutes, in the presence of 1% SDS. As above, samples were run on a 4-20% gradient SDS-polyacrylamide gel, and blotted with 12ca5 monoclonal anti-HA antibodies. The migration of pre-stained molecular weight marker proteins is indicated on the left.


A
Trypsin
-
0'
10'
60'
111
80
61
49
36
25
19
13
B
Trypsin
-
+ 1% SDS
173
111
80
61
49
36
25
19

REFERENCES

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

Achsel, T., Ahrens, K., Brahms, H., Teigelkamp, S., and Luhrmann, R. (1998). The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. Mol Cell Biol *18*, 6756-66.

Aebi, M., Hornig, H., Padgett, R. A., Reiser, J., and Weissmann, C. (1986). Sequence requirements for splicing of higher eukaryotic nuclear pre- mRNA. Cell *47*, 555-65.

Aebi, M., Hornig, H., and Weissmann, C. (1987). 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. Cell *50*, 237-46.

Ban, N., Nissen, P., Hansen, J., Moore, P. B., and Steitz, T. A. (2000). The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 A Resolution. Science *289*, 905-920.

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

Bernstein, H. D. (2000). A surprising function for SRP RNA? [news]. Nat Struct Biol *7*, 179-81.

Bone, R., Silen, J. L., and Agard, D. A. (1989). Structural plasticity broadens the specificity of an engineered protease. Nature *339*, 191-5.

Boudvillain, M., de Lencastre, A., and Pyle, A. M. (2000). A tertiary interaction that links active-site domains to the 5' splice site of a group II intron. Nature *406*, 315-8.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., and Boeke, J. D. (1998). Designer deletion strains derived from Saccharomyces cerevisiae S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast *14*, 115-32.

Brow, D. A., and Guthrie, C. (1988). Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. Nature *334*, 213-8.

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

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

Burge, C. B., Tuschl, T. H., and Sharp, P. A. (1998). Splicing of precursors to mRNAs by the spliceosomes. In RNA World II, R. F. Gesteland, T. R. Cech and J. F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 525-560.

Cech, T. R. (1986). The generality of self-splicing RNA: relationship to nuclear mRNA splicing. Cell *44*, 207-210.

Cech, T. R. (1990). Self-splicing of group I introns. Annu. Rev. Biochem. *59*, 543-68.

Cech, T. R., and Bass, B. L. (1986). Biological catalysis by RNA. Annu Rev Biochem *55*, 599-629.

Chanfreau, G., and Jacquier, A. (1994). Catalytic site components common to both splicing steps of a group II intron. Science *266*, 1383-7.

Chanfreau, G., and Jacquier, A. (1993). Interaction of intronic boundaries is required for the second splicing step efficiency of a group II intron. EMBO J. *12*, 5173-80.

Chanfreau, G., Legrain, P., Dujon, B., and Jacquier, A. (1994). Interaction between the first and last nucleotides of pre-mRNA introns is a determinant of 3' splice site selection in *S. cerevisiae*. Nucleic Acids Res. *22*, 1981-7.

Chang, J. S., and McPheeters, D. S. (2000). Identification of a U2/U6 helix la mutant that influences 3' splice site selection during nuclear pre-mRNA splicing. Rna *6*, 1120-30.

Chiara, M. D., Gozani, O., Bennett, M., Champion-Arnaud, P., Palandjian, L., and Reed, R. (1996). Identification of proteins that interact with exon sequences, splice sites, and the branchpoint sequence during each stage of spliceosome assembly. Mol. Cell. Biol. *16*, 3317-26.

Chiara, M. D., Palandjian, L., Feld Kramer, R., and Reed, R. (1997). Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. EMBO J. *16*, 4746-59.

Chiara, M. D., and Reed, R. (1995). A two-step mechanism for 5' and 3' splice-site pairing. Nature *375*, 510-3.

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

Chua, K., and Reed, R. (1999). The RNA splicing factor hSlu7 is required for correct 3' splice-site choice. Nature *402*, 207-10.

Collins, C. A., and Guthrie, C. (1999). Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. Genes Dev *13*, 1970-82.

Conn, G. L., Draper, D. E., Lattman, E. E., and Gittis, A. G. (1999). Crystal structure of a conserved ribosomal protein-RNA complex. Science *284*, 1171-4.

Deirdre, A., Scadden, J., and Smith, C. W. (1995). Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. EMBO J. *14*, 3236-46.

Dietrich, R. C., Incorvaia, R., and Padgett, R. A. (1997). Terminal intron dinucleotide sequences do not distinguish between U2- and U12-dependent introns. Mol Cell *1*, 151-60.

Dix, I., Russell, C. S., O'Keefe, R. T., Newman, A. J., and Beggs, J. D. (1998). Protein-RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae*. RNA *4*, 1239-50.

Fabrizio, P., and Abelson, J. (1992). Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. Nucleic Acids Res. *20*, 3659-64.

Fabrizio, P., and Abelson, J. (1990). Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. Science *250*, 404-9.

Fabrizio, P., Laggerbauer, B., Lauber, J., Lane, W. S., and Lührmann, R. (1997). An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. EMBO J. *16*, 4092-106.

Favre, A., Saintome, C., Fourrey, J. L., Clivio, P., and Laugaa, P. (1998). Thionucleobases as intrinsic photoaffinity probes of nucleic acid structure and nucleic acid-protein interactions. J Photochem Photobiol B *42*, 109-24.

Fouser, L. A., and Friesen, J. D. (1987). Effects on mRNA splicing of mutations in the 3' region of the Saccharomyces cerevisiae actin intron. Mol. Cell. Biol. *7*, 225-30.

Fouser, L. A., and Friesen, J. D. (1986). Mutations in a yeast intron demonstrate the importance of specific conserved nucleotides for the two stages of nuclear mRNA splicing. Cell *45*, 81-93.

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

Frankel, A. D., and Smith, C. A. (1998). Induced folding in RNA-protein recognition: more than a simple molecular handshake. Cell *92*, 149-51.

Gordon, P. M., Sontheimer, E. J., and Piccirilli, J. A. (2000). Metal ion catalysis during the exon-ligation step of nuclear pre-mRNA splicing: extending the parallels between the spliceosome and group II introns. RNA *6*, 199-205.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell *35*, 849-57.

Gutierrez, A., Yoshimura, T., Fuchikami, Y., Soda, K., and Esaki, N. (1998). A mutant D-amino acid aminotransferase with broad substrate specificity: construction by replacement of the interdomain loop Pro119- Arg120-Pro121 by Gly-Gly-Gly. Protein Eng. *11*, 53-8.

Hanna, M. M. (1989). Photoaffinity cross-linking methods for studying RNA-protein interactions. Methods Enzymol *180*, 383-409.

Harlow, E. (1988). Antibodies : a laboratory manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory).

Hodges, P. E., Jackson, S. P., Brown, J. D., and Beggs, J. D. (1995). Extraordinary sequence conservation of the PRP8 splicing factor. Yeast *11*, 337-42.

Kandels, L. S., and Séraphin, B. (1993). Involvement of U6 snRNA in 5' splice site selection. Science *262*, 2035-9.

Kim, C. H., and Abelson, J. (1996). Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5' splice site. RNA *2*, 995-1010.

Konforti, B. B., and Konarska, M. M. (1994). U4/U5/U6 snRNP recognizes the 5' splice site in the absence of U2 snRNP. Genes Dev. *8*, 1962-73.

Kuhn, A. N., and Brow, D. A. (2000). Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. Genetics *155*, 1667-1682.

Kuhn, A. N., Li, Z., and Brow, D. A. (1999). Splicing factor Prp8 governs U4/U6 RNA unwinding during activation of the spliceosome. Mol Cell *3*, 65-75.

Lesser, C. F., and Guthrie, C. (1993). Mutational analysis of pre-mRNA splicing in *Saccharomyces cerevisiae* using a sensitive new reporter gene, *CUP1*. Genetics *133*, 851-63.

Lesser, C. F., and Guthrie, C. (1993). Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science *262*, 1982-8.

Leung, D. W., Chen, E., and Goeddel, D. V. (1989). A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique *1*, 11-5.

Lossky, M., Anderson, G. J., Jackson, S. P., and Beggs, J. D. (1987). Identification of a yeast snRNP protein and detection of snRNP-snRNP interactions. Cell *51*, 1019-1026.

Lucke, S., Klockner, T., Palfi, Z., Boshart, M., and Bindereif, A. (1997). Trans mRNA splicing in trypanosomes: cloning and analysis of a PRP8-homologous gene from Trypanosoma brucei provides evidence for a U5-analogous RNP. EMBO J. *16*, 4433-40.

Luo, H. R., Moreau, G. A., Levin, N., and Moore, M. J. (1999). The human Prp8 protein a common component U2- and U12-dependent spliceosomes. RNA *5*, 893-908.

Luukkonen, B. G., and Séraphin, B. (1998). Construction of an in vivo-regulated U6 snRNA transcription unit as a tool to study U6 function. RNA *4*, 231-8.

Luukkonen, B. G., and Séraphin, B. (1998). Genetic interaction between U6 snRNA and the first intron nucleotide in *Saccharomyces cerevisiae*. RNA *4*, 167-80.

Luukkonen, B. G., and Séraphin, B. (1997). The role of branchpoint-3' splice site spacing and interaction between intron terminal nucleotides in 3' splice site selection in Saccharomyces cerevisiae. EMBO J. *16*, 779-92.

Mace, J. E., Wilk, B. J., and Agard, D. A. (1995). Functional linkage between the active site of alpha-lytic protease and distant regions of structure: scanning alanine mutagenesis of a surface loop affects activity and substrate specificity. J. Mol. Biol. *251*, 116-34.

MacMillan, A. M., Query, C. C., Allerson, C. R., Chen, S., Verdine, G. L., and Sharp, P. A. (1994). Dynamic association of proteins with the pre-mRNA branch region. Genes Dev. *8*, 3008-20.

Madhani, H. D., Bordonne, R., and Guthrie, C. (1990). Multiple roles for U6 snRNA in the splicing pathway. Genes Dev. *4*, 2264-77.

Madhani, H. D., and Guthrie, C. (1994). Dynamic RNA-RNA interactions in the spliceosome. Annu Rev Genet *28*, 1-26.

Madhani, H. D., and Guthrie, C. (1994). Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome. Genes Dev. *8*, 1071-86.

McPheeters, D. S., Schwer, B., and Muhlenkamp, P. (2000). Interaction of the yeast DExH-box RNA helicase prp22p with the 3' splice site during the second step of nuclear pre-mRNA splicing. Nucleic Acids Res *28*, 1313-21.

Merendino, L., Guth, S., Bilbao, D., Martinez, C., and Valcarcel, J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. Nature *402*, 838-41.

Michel, F., and Ferat, J.-L. (1995). Structure and activities of group II introns. Annu. Rev. Biochem. *64*, 435-61.

Michel, F., and Ferat, J. L. (1995). Structure and activities of group II introns. Annu. Rev. Biochem. *64*, 435-61.

Moore, M. J., and Query, C. C. (2000). Joining of RNAs by splinted ligation. Methods Enzymol *317*, 109-23.

Moore, M. J., Query, C. C., and Sharp, P. A. (1993). Splicing of precursors to mRNA by the spliceosome. In The RNA World, R. F. Gesteland and J. F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 303-57.

Moore, M. J., and Sharp, P. A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature *365*, 364-8.

Moore, M. J., and Sharp, P. A. (1992). Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. Science *256*, 992-7.

Morton, A., and Matthews, B. W. (1995). Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: linkage of dynamics and structural plasticity. Biochemistry *34*, 8576-88.

Muhlrad, D., Hunter, R., and Parker, R. (1992). A rapid method for localized mutagenesis of yeast genes. Yeast *8*, 79-82.

Myers, E. W., and Miller, W. (1988). Optimal alignments in linear space. Comput Appl Biosci *4*, 11-7.

Narlikar, G. J., and Herschlag, D. (1997). Mechanistic aspects of enzymatic catalysis: lessons from comparison of RNA and protein enzymes. Annu Rev Biochem *66*, 19-59.

Newman, A. (1994). RNA splicing. Activity in the spliceosome. Curr. Biol. *4*, 462-4.

Newman, A., and Norman, C. (1991). Mutations in yeast U5 snRNA alter the specificity of 5' splice-site cleavage. Cell *65*, 115-23.

Newman, A. J. (1997). The role of U5 snRNP in pre-mRNA splicing. EMBO J. *16*, 5797-800.

Newman, A. J. (1997). The role of U5 snRNP in pre-mRNA splicing. EMBO J. *16*, 5797-800.

Newman, A. J., and Norman, C. (1992). U5 snRNA interacts with exon sequences at 5' and 3' splice sites. Cell *68*, 743-54.

Nilsen, T. W. (1998). RNA-RNA interactions in nuclear pre-mRNA splicing. In RNA Structure and Function, R. Simons and M. Grunberg-Manago, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 279-307.

Nilsen, T. W. (1994). RNA-RNA interactions in the spliceosome: unraveling the ties that bind. Cell *78*, 1-4.

Nissen, P., Hansen, J., Ban, N., Moore, P. B., and Steitz, T. A. (2000). The Structural Basis of Ribosome Activity in Peptide Bond Synthesis. Science *289*, 920-930.

Noble, S. M., and Guthrie, C. (1996). Transcriptional pulse-chase analysis reveals a role for a novel snRNP-associated protein in the manufacture of spliceosomal snRNPs. Embo Journal *15*, 4368-79.

O'Keefe, R. T., and Newman, A. J. (1998). Functional analysis of the U5 snRNA loop 1 in the second catalytic step of yeast pre-mRNA splicing. EMBO J. *17*, 565-74.

O'Keefe, R. T., Norman, C., and Newman, A. J. (1996). The invariant U5 snRNA loop 1 sequence is dispensable for the first catalytic step of pre-mRNA splicing in yeast. Cell *86*, 679-89.

Padgett, R. A., Podar, M., Boulanger, S. C., and Perlman, P. S. (1994). The stereochemical course of group II intron self-splicing. Science *266*, 1685-8.

Pan, T., Long, D. M., and Uhlenbeck, O. C. (1993). Divalent metal ions in RNA folding and catalysis. In The RNA World, R. F. Gesteland and J. F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 271-301.

Parker, R., and Siliciano, P. G. (1993). Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. Nature *361*, 660-2.

Patterson, B., and Guthrie, C. (1991). A U-rich tract enhances usage of an alternative 3' splice site in yeast. Cell *64*, 181-187.

Pikielny, C. W., and Rosbash, M. (1985). mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. Cell *41*, 119-26.

Pyle, A. M. (1996). Catalytic reaction mechanisms and structural features of Group II intron ribozymes. In Catalytic RNA, F. Eckstein and D. M. J. Lilley, eds. (Berlin ; New York: Springer), pp. 75-107.

Raghunathan, P. L., and Guthrie, C. (1998). RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2. Curr. Biol. *8*, 847-55.

Reed, R., and Maniatis, T. (1985). Intron sequences involved in lariat formation during pre-mRNA splicing. Cell *41*, 95-105.

Reyes, J. L., Gustafson, H. E., Luo, H. R., Moore, M. J., and Konarska, M. M. (1999). The C-terminal region of hPrp8 interacts with the conserved GU dinucleotide at the 5' splice site. RNA *5*, 167-79.

Reyes, J. L., Kois, P., Konforti, B. B., and Konarska, M. M. (1996). The canonical GU dinucleotide at the 5' splice site is recognized by p220 of the U5 snRNP within the spliceosome. RNA *2*, 213-25.

Rothstein, R. (1991). Targetin, Disruption, Replacement, and Allele Rescue: Integrative DNA Transformation in Yeast. In Guide to Yeast Genetics and Molecular Biology, C. Guthrie and G. R. Fink, eds. (San Diego, CA: Academic Press), pp. 281-301.

Ruis, B. L., Kivens, W. J., and Siliciano, P. G. (1994). The interaction between the first and last intron nucleotides in the second step of pre-mRNA splicing is independent of other conserved intron nucleotides. Nucleic Acids Res. *22*, 5190-5.

Ruskin, B., and Green, M. R. (1985). Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. Nature *317*, 732-4.

Schwer, B., and Guthrie, C. (1992). A conformational rearrangement in the spliceosome is dependent on Prp16 and ATP hydrolysis. EMBO J. *11*, 5033-9.

Schwer, B., and Guthrie, C. (1992). A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. Embo J *11*, 5033-9.

Sharp, P. A. (1985). On the origin of RNA splicing and introns. Cell *42*, 397-400.

Sharp, P. A., and Burge, C. B. (1997). Classification of introns: U2-type or U12-type. Cell *91*, 875-9.

Shukla, G. C., and Padgett, R. A. (1999). Conservation of functional features of U6atac and U12 snRNAs between vertebrates and higher plants. RNA *5*, 525-38.

Siatecka, M., Reyes, J. L., and Konarska, M. M. (1999). Functional interactions of Prp8 with both splice sites at the spliceosomal catalytic center. Genes Dev *13*, 1983-93.

Sontheimer, E. J., Gordon, P. M., and Piccirilli, J. A. (1999). Metal ion catalysis during group II intron self-splicing: parallels with the spliceosome. Genes Dev *13*, 1729-41.

Sontheimer, E. J., and Steitz, J. A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. Science *262*, 1989-96.

Sontheimer, E. J., Sun, S., and Piccirilli, J. A. (1997). Metal ion catalysis during splicing of premessenger RNA. Nature *388*, 801-5.

Staley, J. P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell *92*, 315-26.

Staley, J. P., and Guthrie, C. (1999). An RNA switch at the 5' splice site requires ATP and the DEAD box protein Prp28p. Mol Cell *3*, 55-64.

Strauss, E. J., and Guthrie, C. (1991). A cold-sensitive mRNA splicing mutant is a member of the RNA helicase gene family. Genes Dev. *5*, 629-41.

Strobel, S. A. (1999). A chemogenetic approach to RNA function/structure analysis. Curr Opin Struct Biol *9*, 346-52.

Szewczak, A. A., Ortoleva-Donnelly, L., Ryder, S. P., Moncoeur, E., and Strobel, S. A. (1998). A minor groove RNA triple helix within the catalytic core of a group I intron. Nat Struct Biol *5*, 1037-42.

Tarn, W. Y., and Steitz, J. A. (1996). Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns. Science *273*, 1824-32.

Teigelkamp, S., Newman, A. J., and Beggs, J. D. (1995). Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. EMBO J. *14*, 2602-12.

Umen, J. G., and Guthrie, C. (1996). Mutagenesis of the yeast gene *PRP8* reveals domains governing the specificity and fidelity of 3' splice site selection. Genetics *143*, 723-39.

Umen, J. G., and Guthrie, C. (1995). A novel role for a U5 snRNP protein in 3' splice site selection. Genes Dev. *9*, 855-68.

Umen, J. G., and Guthrie, C. (1995). Prp16p, Slu7p, and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. RNA *1*, 584-97.

Umen, J. G., and Guthrie, C. (1995). The second catalytic step of pre-mRNA splicing. RNA *1*, 869-85.

Valadkhan, S., and Manley, J. L. (2000). A tertiary interaction detected in a human U2-U6 snRNA complex assembled in vitro resembles a genetically proven interaction in yeast. RNA *6*, 206-19.

Vidal, V. P., Verdone, L., Mayes, A. E., and Beggs, J. D. (1999). Characterization of U6 snRNA-protein interactions. RNA *5*, 1470-81.

Vijayraghavan, U., Parker, R., Tamm, J., Iimura, Y., Rossi, J., Abelson, J., and Guthrie, C. (1986). Mutations in conserved intron sequences affect multiple steps in the yeast splicing pathway, particularly assembly of the spliceosome. EMBO J. *5*, 1683-95.

Walberer, B. J. (2000). Construction and analysis of a complete database of hydrogen-bonded base combinations. In Department of Biophysics and Biochemistry (San Francisco: University of California, San Francisco).

Wassarman, D. A., and Steitz, J. A. (1992). Interactions of small nuclear RNA's with precursor messenger RNA during in vitro splicing. Science *257*, 1918-25.

Weeks, K. M. (1997). Protein-facilitated RNA folding. Curr Opin Struct Biol *7*, 336-42.

Weeks, K. M., and Cech, T. R. (1996). Assembly of a ribonucleoprotein catalyst by tertiary structure capture. Science *271*, 345-8.

Weeks, K. M., and Cech, T. R. (1995). Efficient protein-facilitated splicing of the yeast mitochondrial bI5 intron. Biochemistry *34*, 7728-38.

Weeks, K. M., and Cech, T. R. (1995). Protein facilitation of group I intron splicing by assembly of the catalytic core and the 5' splice site domain. Cell *82*, 221-30.

Weiner, A. M. (1993). mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? Cell *72*, 161-4.

Wimberly, B. T., Guymon, R., McCutcheon, J. P., White, S. W., and Ramakrishnan, V. (1999). A detailed view of a ribosomal active site: the structure of the L11- RNA complex. Cell *97*, 491-502.

Wu, S., Romfo, C. M., Nilsen, T. W., and Green, M. R. (1999). Functional recognition of the 3' splice site AG by the splicing factor U2AF35. Nature *402*, 832-5.

Wyatt, J. R., Sontheimer, E. J., and Steitz, J. A. (1992). Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. Genes Dev. *6*, 2542-53.

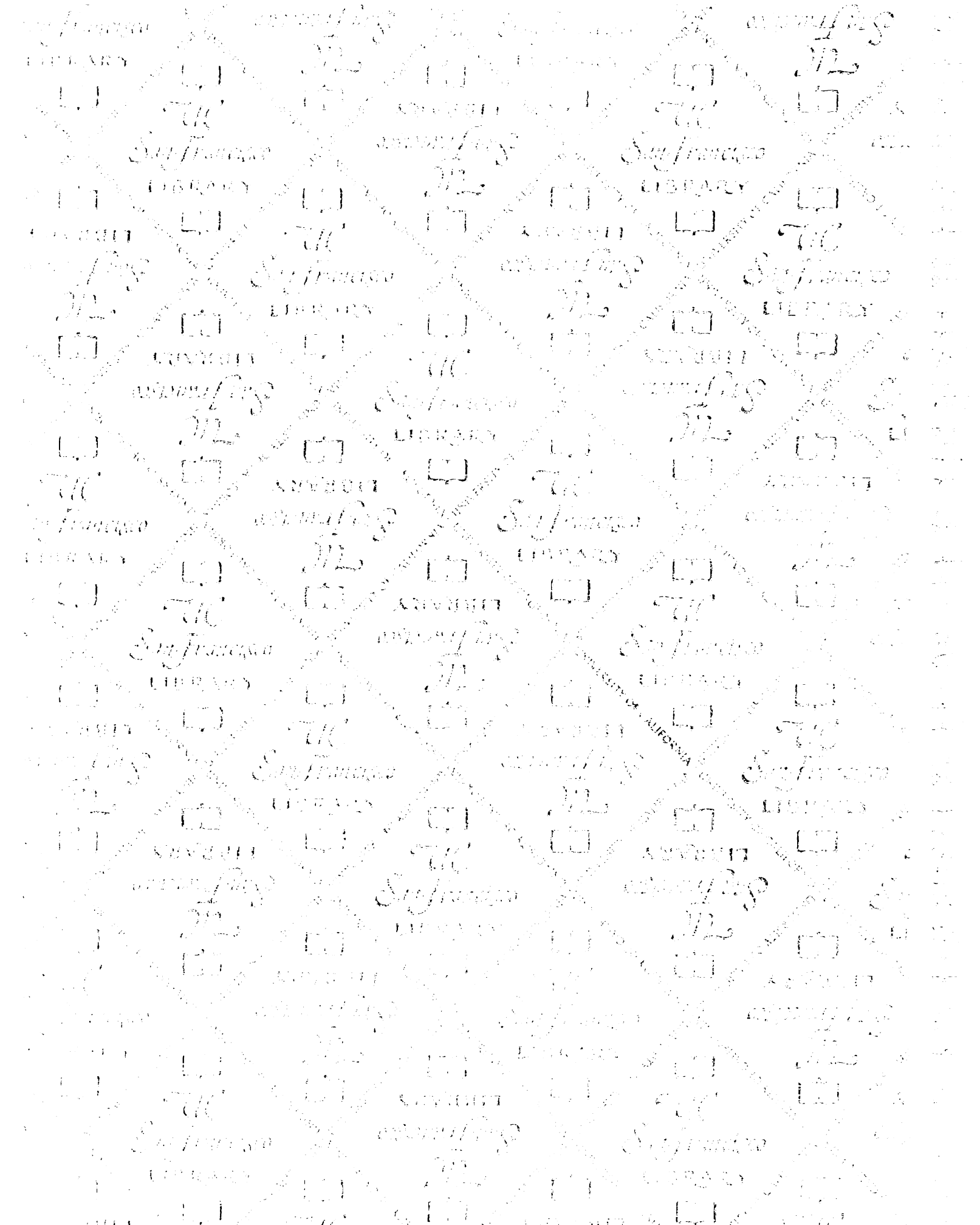
Yarus, M. (1993). How many catalytic RNAs? Ions and the Cheshire cat conjecture. Faseb J *7*, 31-9.

Yean, S.-H., Wuenschell, G., Termini, J. and Lin, R.-J. (2000). Metal ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature *408*, 881-884.

Yu, Y.-T., Scharl, E. C., Smith, C. M., and Steitz, J. A. (1998). The growing world of small nuclear ribonucleoproteins. In RNA World II, R. F. Gesteland, T. R. Cech and J. F. Atkins, eds. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press), pp. 487-524.

Yu, Y. T., Maroney, P. A., Darzynkiwicz, E., and Nilsen, T. W. (1995). U6 snRNA function in nuclear pre-mRNA splicing: a phosphorothioate interference analysis of the U6 phosphate backbone. RNA *1*, 46-54.

Zorio, D. A., and Blumenthal, T. (1999). Both subunits of U2AF recognize the 3' splice site in Caenorhabditis elegans. Nature *402*, 835-8.


THE
LIBRARY
UNIVERSITY OF CALIFORNIA