Title
The role of protein-RNA interactions in 3' splice site selection during pre-mRNA splicing

Author
Umen, James G.

Publication Date
1995

Peer reviewed|Thesis/dissertation

The Role of Protein-RNA Interactions in 3' Splice Site Selection During pre-mRNA Splicing

by

James G. Umen

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO — SANTA BARBARA • SANTA CRUZ

CHRISTINE GUTHRIE
SCHOOL OF MEDICINE
DEPARTMENT OF BIOCHEMISTRY AND BIOPHYSICS

SAN FRANCISCO, CALIFORNIA 94143-0448
TELEPHONE: (415) 476 2321
TELEFAX: (415) 502 5306

September 22, 1995

Genes and Development
Cold Spring Harbor Laboratory Press
Box 100
1 Bungtown Road
Cold Spring Harbor, New York 11724

To whom it may concern,

I would like the permission of Genes and Development to reprint, for my Ph.D. thesis the article "A Novel Role for a U5 snRNP Protein in 3' Splice Site Selection" by myself and Christine Guthrie (Genes and Development, 9: 855-868). The dissertation will be placed on microfilm by University Microfilms who request permission to supply single copies on demand. Please respond by fax to 415-502-5306. A timely reply will be greatly appreciated.

Thank you for your consideration.

Sincerely,

James G. Umen

Permission granted by the copyright owner, contingent upon the consent of the original author, provided complete credit is given to the original source and copyright date.

BY _____ DATE 10/10/95

Cold Spring Harbor Laboratory, POB 100, Cold Spring Harbor, NY 11724

TOTAL P.01

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

CHRISTINE GUTHRIE
SCHOOL OF MEDICINE
DEPARTMENT OF BIOCHEMISTRY AND BIOPHYSICS

SAN FRANCISCO, CALIFORNIA 94143-0448
TELEPHONE: (415) 476 2321
TELEFAX: (415) 502 5306

September 22, 1995

RNA
Publication Office
Cambridge University Press
40 West 20th Street
New York, New York 10011-4211

To whom it may concern,

I would like the permission of RNA to reprint, for my Ph.D. thesis the article "Prp16, Slu7 and Prp8 interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing" by myself and Christine Guthrie (RNA, Manuscript # 95-083, in press). The dissertation will be placed on microfilm by University Microfilms who request permission to supply single copies on demand. Please respond by fax to 415-502-5306. A timely reply will be greatly appreciated.

Thank you for your consideration.

Sincerely,

James G. Umen

RNA SOCIETY

9650 Rockville Pike
Bethesda, Maryland 20814-3998

Telephone: 301-530-7120 FAX: 301-530-7049 E-mail: rna@faseb.org

October 4, 1995

Mr. James G. Umen
Dept. Biochemistry and Biophysics
School of Medicine
University of California
San Francisco, CA 94143-0448

Dear Mr. Umen,

As owner of the copyright in the journal RNA, the RNA Society hereby grants your request for use of the materials as described in your attached letter provided that such use includes full citation to the corresponding article in RNA.

Sincerely,

Chris L. Greer
CEO, RNA Society

CLG/hs

att: Ltr from James G. Umen

ACKNOWLEDGMENTS

Although it is wonderful to finally be finished with my thesis, I always knew that I would feel sad when I wrote this section. At times graduate school has seemed endless, but I already miss it (at least some of it). First and foremost, my deepest gratitude is owed to my advisor, Christine, who had to wait six years for me to finally figure out that it wouldn't hurt to agree with her occasionally. In fact, it has been her uncanny ability to see things from a completely different perspective than me, and frequently sway me to her viewpoint, that has so enriched my graduate training. Christine's insight into scientific problems and her beautifully precise ability to express herself are continuing sources of inspiration.

The entire Guthrie lab deserves thanks, not just for tolerating me, since they had little choice, but for tolerating me so good naturedly. Many people, past and present, have contributed to making my experiences in the lab interesting, enjoyable and scientifically stimulating. When I was a roton, Paul Siliciano, got me hopelessly entangled with U1 snRNA, from which I have only just now extricated myself by writing a lengthy thesis appendix. Despite this, he was a great rotation supervisor and helped me get started in the RNA world. My first baymate, Sean Burgess, was an ideal and understanding brooding companion for the dark years of graduate school. She also reminded me to savor the sunsets from our perfect 9th floor vantage point. Yan Wang (Wang Yan, actually) took Sean's place in Bay B. Although she carefully avoided becoming my Chinese teacher, she could not have been a more considerate baymate (with the exception of being able to outclone me within one month of learning to do a ligation). Since moving to the 2nd floor, I have had the fortune to end up in the bay with the highest concentration of Y chromosomes. Leslie Stewart, deserves special commendation for seeming to tolerate this situation. She has also helped diffuse the reputation of autocratic lunacy associated with the job of radiation czar, previously held by me and Anita Jandrositz. Y chromosome #1 in my bay, Chris Siebel, has done much

to help me both within and outside the realm of science. Being a pessimist, it continually amazes me to see his optimism and determination in pursuing scientific goals. He also introduced me to the thrills of abalone diving and surfing, making it that much harder to leave California. Ychromosome #2 in the bay is Jon Staley, the person who has to keep teaching me the chemistry which I would otherwise fabricate. Jon and Chris have also taught me patience. Even though they are slow talkers, I found that it pays to keep my mouth shut and listen rather than finishing their sentences for them. Cammie Lesser gets full credit for introducing me to the ecstasy and misery of copper genetics, both of which contributed immensely to my graduate training. Dan Frank paved the way for the genetic analysis of second step factors which dominated much of my work, consumed much of my time, and has defied depiction in a non-boy-like fashion. My classmate and labmate Hiten Madhani was a chief source of interesting scientific discourse and useful criticism througout most of my graduate career. Despite his intimidating and encyclopedic knowledge of many scientific subjects (+ the solar system) I can still take some small satisfaction in being able to beat him at Scrabble. Evi Strauss provided a great deal of creative inspiration to me and the lab. I will never forget all the time we spent writing goofy lyrics to lab musical videos. Shelly Jones has done many wonderful things for me and the lab over the years. These include starting up Second Step Club again, giving up on journal club (eventually), and being the one person in lab who expresses as many opinions as me. Tanya Awabdy has done an amazing job of taking over many of Shelly's unofficial lab responsibilities, including reminding me of how many holes are in my green sweater. Anne de Bruyn Kops, the best talker in the lab, always told me interesting stories about two of my favorites topics, food and travel. She also has the distinction of possessing a weirder and more difficult to pronounce last name than me. Although she has given up on Prp8, Amy Kistler is still the most enthusiastic lab member. If she had only joined the lab earlier, she might have turned me and Ping into rock climbers. My three rotons, Louise Chang, Tom Wang and Cathy Collins probably taught me much

more than I was able to teach them. Cathy has already done amazingly well despite all my advice, and will no doubt carry Prp8 to its rightful position of supreme dominance among splicing factors. Pratima Raghunathan has all of the reactionary argumentativeness that is required to take over my job as lab critic. This skill has been demonstrated to me (on me?) on numerous occasions. If she can just get over being so nice and soft spoken most of the time, she will, no doubt, surpass me in this lab job. Suzanne Noble is the only lab member whose mother has almost given me a vicarious ulcer. But that is mostly Suzanne's fault. Suzanne herself has always taken great pleasure in pointing out all of my oddities and eccentricities. The best part about being analyzed this way is that all of her own eccentricities are illuminated in the process. In addition Suzanne has given me hours of arguing pleasure over many inconsequential topics. Both she and Pratima have kept me on my toes and prevented me from getting away with most of my outlandish statements. Kent Duncan (replacement Y chromosome) and Maki Inada, the next generation of surfers and splicers, will be lucky enough to get through graduate school with only a fraction of the unsolicited opinions from me that others have had to endure and combat. Lucita Esperas will always hold a special place in my heart and stomach. Besides being the best plate and gel pourer in the world, she has supplied me with a continuous source of topical commentary on two shared causes in our lives -- food and demanding overseas relatives. Heli Roiha has played an instrumental role in getting me through graduate school with no scientific needs left unattended. She has also reintroduced me to the English language from an utterly different perspective. "Riding the green wave" is far more interesting than driving on a street with synchronized lights. Carol Pudlow is another person who has introduced me to a new perspective on the world. Should I ever contemplate a career in infomercials, I know whom to call.

I thank my thesis committee members, Ira Herskowitz and Peter Walter, for their continual confidence in my ability to get through graduate school and produce this

document. They provided useful commentary and advice at critical times in my graduate career. Ira is also responsible for giving me a crash course in certain aspects of genetics and grammar. This lesson occured in the aftermath of what has to be one of the most amazingly fast and thorough thesis readings in dissertation history.

Finally, I thank my parents and entire family for their love and generous support throughout my life and graduate career. My wife, Qingping Li, gets special thanks for being such a wonderful and supportive companion and friend during my time in graduate school. Qingping's eccentricities are a perfect complement to my own, and she has helped me maintain the balance and perspective required to produce a 200^{+}-page thesis while retaining a moderate level of sanity.

THE ROLE OF PROTEIN-RNA INTERACTIONS IN 3' SPLICE SITE SELECTION DURING pre-mRNA SPLICING

James G. Umen

ABSTRACT

Removal of introns from pre-mRNAs requires accurate identification of consensus sequences at the 5' splice site, branchsite, and 3' splice site. Two elements contribute to 3' splice site identification in yeast, an invariant PyAG motif and an upstream uridine-rich tract. In this work, an *in vivo* 3' splice site cis-competition assay was used to select yeast mutants that are impaired for recognizing the conserved uridine-rich tract preceding the 3' splice junction. A novel allele of a U5 snRNP protein, Prp8, was identified in this selection. Besides impairing uridine tract recognition, the *prp8-101* mutation exacerbates the defects due to alterations in the PyAG motif at the 3' splice junction. *prp8-101* also displays specific genetic interactions with mutant alleles of genes encoding step two-specific splicing proteins. Site-specific crosslinking reveals that Prp8 mediates 3' splice site selection through direct binding to the 3' splice site, and the *prp8-101* mutation weakens this binding. Prp16, an RNA-dependent ATPase, and Slu7 also crosslink to the 3' splice site. These two proteins are required for the second catalytic step of splicing, and Slu7 is also implicated in 3' splice site selection. Binding of these factors to the 3' splice site occurs in two stages, with Prp16 binding first. After hydrolysis of ATP, Prp16 releases the 3' splice site and this allows Prp8 and Slu7 to bind. This two-step binding regime may reflect a mechanism for enhancing the fidelity of 3' splice site recognition. The kinetics of 3' splice site crosslinking suggest that Prp8 may be at or near the active site during the second catalytic step. In support of this idea, a novel class of *PRP8* alleles that can suppress the effects of PyAG alterations was isolated. The complex pattern of preferences for different PyAG mutations exhibited by these

alleles suggests that Prp8 directly contacts these nucleotides and/or the spliceosomal active site during the second catalytic step. Thus, Prp8 contributes to the specificity of splice site choice and to the fidelity of 3' splice site utilization.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Intron Recognition Problem

One of the most fascinating discoveries in the recent era of molecular biology has been that of split genes in eukaryotes (Berget et al., 1977; Chow et al., 1977). The unexpected finding that the functional coding sequences for nuclear genes (exons) are interrupted by nonfunctional spacer sequences (introns) raises two types of questions. First, why are there introns and where did they come from? Second, how are introns removed during the process of gene expression? While the first question is still a great mystery of molecular evolution, the answers to the second question have been more forthcoming and are the subject of intense study. However, the mechanism of splicing has proven to be far more complicated than anyone might have anticipated, thus raising interesting questions about the origins and design principles of the splicing machinery.

The chemical pathway of intron excision involves two successive transesterification reactions (reviewed in Guthrie, 1991; Figure 1). In the first step, the phosphodiester bond at the upstream exon/intron junction (5' splice site) is attacked by the 2' hydroxyl group of an adenosine residue in the intron (branchsite) to form an unusual lariat intermediate structure and a free upstream exon. In the second step, the phosphodiester bond at the downstream intron/exon junction (3' splice site) is attacked by the free 3' hydroxyl group of the first exon to form ligated exons and an excised lariat intron.

Introns are identified by the presence of conserved sequences at the 5' splice site, branch site, and 3' splice site. These sequences are recognized by a set of small ribonucleoprotein particles (U1, U2, U4, U5 and U6 snRNPs) which, along with accessory proteins, assemble onto pre-mRNAs and form the machine which catalyzes splicing, the spliceosome (Figure 2; reviewed in Green, 1991; Guthrie, 1991; Rymond & Rosbash, 1992; Moore et al., 1993). One of the key properties of the spliceosome is its ability to identify proper splice sites accurately. In metazoans, where genes often contain dozens of introns, an error rate of just 1% in splice site choice would lead to a substantial fraction of improperly spliced messages encoding truncated, aberrant, and possibly deleterious proteins. Thus, one might predict that the mechanism of splicing will entail

multiple recognition events to ensure proper splice site choice. Indeed it has now been established that the 5' splice site is recognized at least twice by three different snRNPs. During the course of spliceosome assembly, the 5' splice site is first bound by U1, which is later displaced by U5 and U6 binding. U2 is responsible for branch site recognition and appears to remain stably associated with this sequence throughout the splicing reaction (see above for citations).

The subject of this thesis is the mechanism of 3' splice site selection. When this work was started, nothing was known about the spliceosomal components involved in 3' splice site recognition. Without any obvious base pairing partners among the snRNAs, it was not clear how 3' splice site recognition factors might be identified; nor was it clear whether such factors would be RNAs, proteins or a combination of the two. Numerous proteins had been found to bind the 3' splice site in mammalian extracts (Chabot et al., 1985; Gerke & Steitz, 1986; Tazi et al., 1986; Swanson & Dreyfuss, 1988; Garcia-Blanco et al., 1989; Buvoli et al., 1990), but with the exception of U2AF, a protein required for branchsite binding by U2 (Ruskin et al., 1988), their functional significance was not known. What had been established is that the 3' splice site is recognized through at least two sequences. The first is an invariant PyAG motif at the 3' splice junction (Reed & Maniatis, 1985; Ruskin & Green, 1985; Vijayraghavan et al., 1986; Fouser & Friesen, 1987), and the second is an upstream pyrimidine-rich tract (Reed, 1989; Patterson & Guthrie, 1991) (Figure 2). The assay that was developed to determine the function of the pyrimidine-rich tract in 3' splice site selection in *Saccharomyces cerevisiae* involved tandem 3' splices sites competing in cis. It was found that a uridine-rich tract placed upstream of a 3' splice would enhance usage of that splice site at the expense of its uridine-poor competitor (Patterson & Guthrie, 1991).

In Chapter 1, which is reprinted from a paper published in *Genes & Development*, I used this type of 3' splice site cis-competition construct as a genetic tool for identifying 3' splice site recognition factors. We reasoned that there was a splicing factor which

recognized the uridine-rich tract and activated use of an adjacent 3' splice site. Loss of uridine recognition function by this factor was predicted to cause the uridine-rich 3' splice site to be less strongly favored and thereby activate use of its uridine-poor competitor. By fusing the uridine-poor 3' splice site in frame with a selectable marker, *CUP1*, I was able to select for uridine-tract recognition mutants. In doing so, I found a novel allele of a previously characterized splicing factor, *PRP8*.

Prp8 is a large (280 kd), highly conserved U5 snRNP protein whose function in the spliceosome was not known (Anderson et al., 1989; Pinto & Steitz, 1989; Hodges et al., 1995). It was particularly interesting to find a U5 snRNP protein involved in 3' splice site selection since a loop structure in U5 snRNA had recently been found to interact genetically with exon sequences adjacent to both 5' and 3' splice sites (Newman & Norman, 1991; Newman & Norman, 1992; Figure 2) and to crosslink to those exon sequences in vitro (Sontheimer & Steitz, 1993). I utilized a site-specific crosslinking assay to demonstrate that Prp8 directly contacts the 3' splice site during the splicing reaction. Thus, the role of Prp8 in 3' splice site selection appears to be direct, and it may act in concert with U5 snRNA to facilitate exon alignment during the second step of splicing.

Chapter 2, which is reprinted from a paper published in *RNA*, is a more detailed biochemical examination of 3' splice site binding factors. First, I show that the novel allele of *PRP8* which I isolated, *prp8-101*, causes a 3' splice site binding defect in Prp8. *prp8-101* was also found to interact genetically with alleles of genes that encode proteins specific for the second catalytic step of splicing, *PRP16*, *PRP17*, *PRP18* and *SLU7*. Two of these proteins, Prp16 and Slu7 can also be crosslinked to the 3' splice site. Beate Schwer, when she was a postdoctoral fellow in the lab, had shown that Prp16 is an RNA-dependent ATPase that elicits a conformational change in the spliceosome (Schwer & Guthrie, 1992), and Dan Frank, a former graduate student, had shown that Slu7 is involved in 3' splice site selection (Frank & Guthrie, 1992). By analyzing the 3' splice site binding kinetics of Slu7, Prp16 and Prp8, I was able to show that the 3' splice site is

recognized in two distinct stages. In the first stage, Prp16 binding is strong and Slu7/Prp8 binding is weak. In the second stage, which requires ATP hydrolysis by Prp16 and the functions of Prp17 and Prp18, Slu7 and Prp8 bind strongly to the 3' splice site and Prp16 binding weakens. This two-stage recognition process may be involved in maintaining the fidelity of 3' splice site selection. Furthermore, the kinetics of 3' splice site binding by these factors suggests that Prp8 and Slu7 are at, or near, the spliceosomal active site during catalysis.

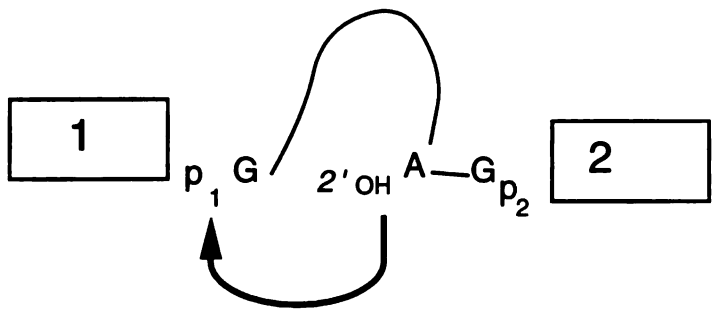
The work in Chapter 3, which has been submitted for publication in *Genetics*, was motivated by the possibility of finding a functional role for Prp8 at the spliceosomal active site. Furthermore, it was of great interest to begin assigning specific functions to different structural domains of this massive protein. To these ends, I constructed a mutagenized *PRP8* library from which two classes of mutants were selected. The first were uridine recognition mutants similar to *prp8-101*. The mutations that generate this phenotype, including the original *prp8-101* allele, were found to affect one of two amino acids in the C-terminal portion of the protein.

A second type of *PRP8* mutation was also found in this study. This class of mutant is capable of suppressing the effects of alterations in the PyAG motif at the 3' splice junction. These alleles are functionally distinct from the *prp8-101*-like alleles in two ways. First, they suppress the effects of PyAG alterations, rather than exacerbating them. Second, they do not affect uridine tract recognition. Interestingly, the PyAG suppressor alleles map to a region just upstream of the *prp8-101*-like class. The complex spectrum of preferences for different PyAG alterations displayed by these alleles of *PRP8* strongly suggests a direct interaction between Prp8 and the PyAG trinucleotide at the spliceosomal active site. Thus, Prp8 has a role in specifying the location of a proper 3' splice site and in maintaining the fidelity of the 3' splice site recognition process.

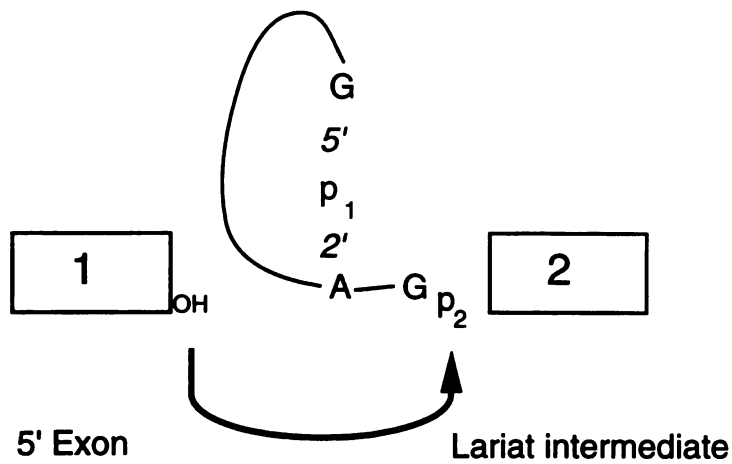
In the epilogue, I integrate my work on proteins involved in 3' splice site selection with the work of others in a comprehensive review on the second catalytic step of pre-

mRNA splicing. In the course of my thesis work, important progress has been made in understanding 3' splice site recognition and the second catalytic step. It is particularly exciting that we can now begin to construct an RNA-based model of the step two active site and are not far from incorporating proteins into the picture. Thus, over a period of six years, our perspective on the problem of 3' splice site selection has taken a quantum leap forward, from a virtual information vacuum to a detailed (albeit still incomplete) framework from which explicit mechanistic models can now be formulated.

Figure 1. The two chemical steps of splicing. Depicted are exons 1 and 2, phosphate and guanosine residues at 5' and 3' splice sites, and adenosine residue at the branchsite.



Step 1 ↓ **5' Splice Site Cleavage and Branch Formation**



Step 2 ↓ **3' Splice Site Cleavage and Exon Ligation**

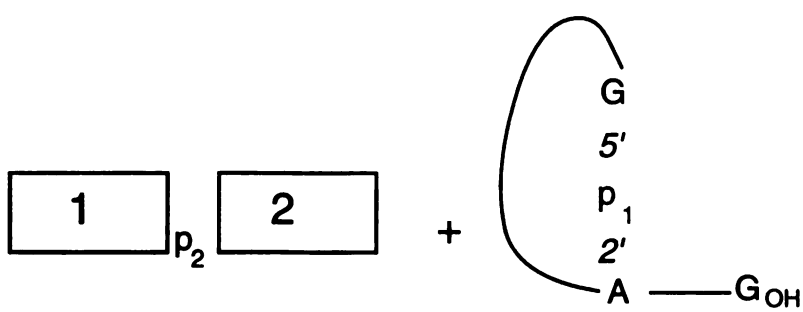
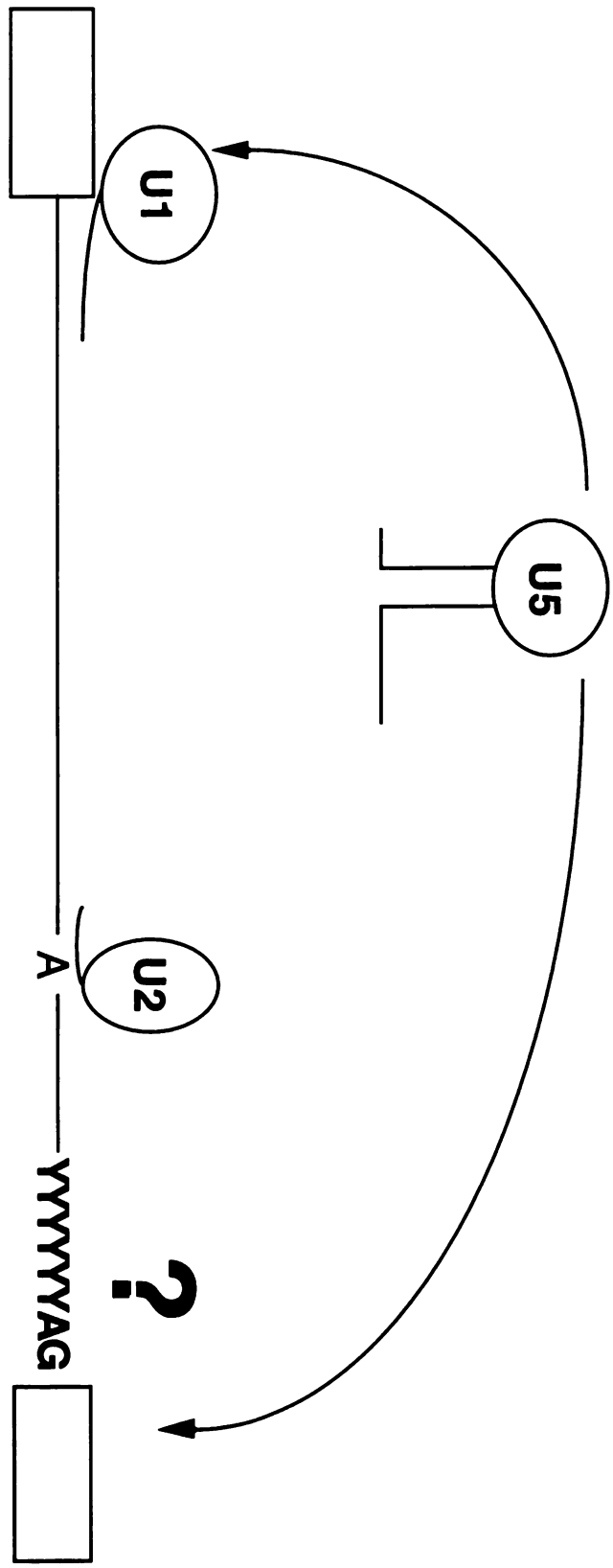


Figure 2. Intron recognition in yeast. Depicted is an intron with U1 and U2 snRNPs bound to the 5' splice site and branchsite. U5 interacts with exon sequences flanking the splice sites. The conserved nucleotides at the 3' splice site are shown. The question mark represents the lack of identified factors that recognize the 3' splice site.


U1
U5
U2
A
YYYYYAG
?

CHAPTER 1

A Novel Role for a U5 snRNP Protein in 3' Splice Site Selection

Abstract

The choice of a 3' splice site in *Saccharomyces cerevisiae* introns involves recognition of a uridine-rich tract upstream of the AG dinucleotide splice junction. By isolating mutants that eliminate the normal preference for uridine-containing 3' splice sites in a cis-competition, we identified a mutation which is an allele of *PRP8*, *prp8-101*. This was unexpected since previous analysis has demonstrated that the U5 snRNP protein encoded by *PRP8* is required for spliceosome assembly prior to the first catalytic step of splicing. In contrast, the uridine recognition defect caused by the *prp8-101* mutation selectively inhibits the second catalytic step of splicing. This defect is seen not only in 3' splice site cis-competitions, but also in the splicing of an unusual intron in the *TUB3* gene, and in the *ACT1* intron when utilization of its 3' splice site is rate-limiting for splicing. Consistent with a direct role in 3' splice site selection, Prp8 can be crosslinked to the 3' splice site during the splicing reaction. These data demonstrate a novel function for Prp8 in 3' splice site recognition and utilization.

Introduction

Nuclear pre-mRNA introns are identified by the presence of conserved sequences at their 5' splice site, internal branchsite and 3' splice site. These sequences are recognized by a set of small ribonucleoprotein particles (U1, U2, U4, U5 and U6 snRNA-containing snRNPs) which, along with other protein factors, bind to pre-mRNAs in an ordered set of events that lead to spliceosome assembly and catalysis. The two-step chemical reaction of splicing proceeds via successive transesterifications. First, the 2' hydroxyl of the branchsite adenosine attacks the 5' splice site phosphodiester bond to generate the lariat intermediate and the free 5' exon; next, the 3' hydroxyl of the 5' exon attacks the 3' splice site phosphodiester bond forming ligated exons and an excised lariat intron (reviewed in Green 1991; Guthrie 1991; Rymond and Rosbash 1992; Moore et al. 1993).

Prior to the catalytic steps of splicing, the reaction partners must be accurately selected and juxtaposed. Multiple snRNA-pre-mRNA and snRNA-snRNA interactions are involved in recognizing and aligning the 5' splice site and branchsite before the first catalytic step. In contrast, much less is known about the events that lead to 3' splice site selection and the second catalytic step of splicing. An AG dinucleotide preceded by a pyrimidine-rich tract comprises the 3' splice site of both yeast and mammalian introns. Moreover, an interaction between the first and last nucleotides of yeast introns plays a functional role in 3' splice site utilization (Parker & Siliciano, 1993). However, trans-acting factors that contribute to 3' splice site recognition have been difficult to identify. Two conserved loop nucleotides of U5 snRNA interact with residues in the second exon, although the lack of sequence conservation in exons makes this interaction unlikely to play a major instructive role in 3' splice site selection (Newman & Norman, 1992; Sontheimer & Steitz, 1993). In mammals, several proteins have been shown to bind the 3' splice site (IBP, 70K, U2AF, hnRNP C, hnRNP D, PTB, hnRNP A1, PSF)(Gerke & Steitz, 1986; Tazi et al., 1986; Ruskin et al., 1988; Swanson & Dreyfuss, 1988; Garcia-Blanco et al., 1989; Buvoli et al., 1990; Patton et al., 1993). U2AF and PSF are required for the first and second steps of splicing, respectively (Zamore & Green, 1991; Gozani et al., 1994), but the specific roles that these proteins play in splice site selection have not been determined. Multiple interactions with the 3' splice site probably reflect that this sequence is often required for early steps of spliceosome assembly in

mammals and is also required for later steps in the splicing reaction (Reed & Maniatis, 1985; Ruskin & Green, 1985; Lamond et al., 1987; Reed, 1989; Zhuang & Weiner, 1990).

In the yeast *Saccharomyces cerevisiae*, the uridine-rich tract and AG dinucleotide splice acceptor appear to be required only during the second catalytic step (Rymond & Rosbash, 1985; Rymond et al., 1987). Although the uridine residues in yeast 3' splice sites are generally less conserved than the pyrimidines in mammalian introns (Parker & Patterson, 1987), they play an important role in efficient 3' splice site utilization. Patterson and Guthrie (1991) demonstrated that a uridine-rich tract placed upstream of a 3' splice site greatly enhances that splice site's ability to compete in cis with a non-uridine containing 3' splice site. This, in turn, implies the existence of a trans-acting factor(s) that recognizes this sequence and preferentially recruits uridine-containing 3' splice sites into the spliceosome.

U5 snRNP has been linked genetically and biochemically to the 3' splice site and the second catalytic step of splicing (Chabot et al., 1985; Gerke & Steitz, 1986; Tazi et al., 1986; Patterson & Guthrie, 1987; Winkelmann et al., 1989; Frank et al., 1992; Newman & Norman, 1992; Sontheimer & Steitz, 1993). Therefore, Frank and Guthrie (1992) tested the hypothesis that *slu* mutants, which were identified by their synthetic lethality with mutant U5 snRNAs, are defective in uridine-specific 3' splice site selection. This work uncovered a role for the Slu7 protein in selection of 3' splice sites that are located distally from the branchsite; however, this factor is unlikely to interact directly with the uridine-rich tract (Frank & Guthrie, 1992).

Here we have modified the 3' splice site cis-competition reporters developed by Patterson and Guthrie (1991) for use in a genetic selection for factors that mediate uridine-rich tract recognition in yeast 3' splice sites. We have identified *prp8-101*, a novel mutation in the integral U5 snRNP protein encoded by *PRP8* (Lossky et al., 1987; Jackson et al., 1988); *prp8-101* specifically impairs recognition of the uridine-rich tract in a 3' splice site cis-competition. This finding was unexpected, since all previous alleles of *PRP8* are defective at or prior to the first catalytic step of splicing (Jackson et al., 1988; Brown & Beggs, 1992). The *prp8-101* mutation severely inhibits the second step of splicing of the *TUB3* intron, which depends critically on its uridine tract for efficient 3' splice site recognition. Other introns also require the uridine recognition function of Prp8 when utilization of the 3' splice site is made rate-limiting for splicing. Consistent with a direct role in 3' splice site recognition, we have used site-specific UV

crosslinking to demonstrate that Prp8 contacts the 3' splice site during the splicing reaction. In conjunction with data implicating Prp8 and the U5 snRNP in the first step of splicing (Brown & Beggs, 1992; Wyatt et al., 1992), these results suggest a model in which Prp8 can function to juxtapose the two reaction partners for the second catalytic step.

Results

Isolation of Mutants Defective in 3' Splice Site Uridine Recognition

In order to identify factors involved in uridine tract recognition, we modified two 3' splice site cis-competition constructs developed by Patterson and Guthrie (1991). These constructs each contain a single 5' splice site and branchsite and two potential 3' splice acceptors. In +T PyDOWN (Figure 1A), use of the uridine-rich branchsite-proximal 3' splice site is favored over the branchsite-distal, adenosine-rich 3' splice site (Figure 2A; Table 1). We reasoned that if the factor(s) that recognizes the uridine-rich tract (Polyuridine Splicing Factor or PSF) was mutated to abrogate uridine recognition, then the splicing pattern of +T PyDOWN would be altered to give a more balanced usage of the two competing splice sites (Figure 1B). The net result would be a decrease in splicing to the proximal, uridine-rich 3' splice site and an increase in splicing to the distal, adenosine-rich 3' splice site.

To allow a genetic selection for *psf* mutants, we fused the +T PyDOWN competition to the selectable marker *CUP1*, which encodes the yeast metallothionein homologue and confers dosage dependent copper resistance to cells (Lesser & Guthrie, 1993). In this fusion, use of the branchsite-distal 3' splice site results in a message with the initiator AUG in frame with the *CUP1* coding sequence, and use of the branchsite-proximal 3' splice site results in a message with the *CUP1* coding sequence out of frame. *psf* mutants were selected by their predicted phenotype of increased use of the distal splice site and therefore increased copper resistance (Figure 1B, Methods).

We isolated two independent and allelic mutants, *psf1-1* and *psf1-2*, that displayed similar phenotypes. Primer extension analysis of the +T PyDOWN RNA products from *psf1-1* and its wild type parent strain (Figure 2A, Table 1) revealed a dramatic increase in use of the distal (uridine-poor) 3' splice site and a decrease in use of the proximal (uridine-rich) 3' splice site in the

psf1-1 mutant (proximal/distal = 1.6) versus wild type (proximal/distal = 22). The change in ratio of splice site usage in the *psf1-1* mutant is accompanied by a greater than three fold increase in copper resistance (.18 mM) versus wild type (.05 mM).

In principle, the change in splice site choice in the *psf1-1* strain could result from either a loss of uridine recognition or a change in preference from branchsite-proximal to branchsite-distal 3' splice sites ("polarity"). To distinguish these hypotheses, we used the construct +A WT (Figure 1A) which is the opposite of the +T PyDOWN construct in that the branchsite-proximal 3' splice site is uridine-poor and the distal 3' splice site is uridine-rich. Since branchsite proximity is preferred in a wild type strain (Patterson & Guthrie, 1991), the proximal 3' splice site of +A WT is utilized almost as efficiently as the distal 3' splice site. We introduced this construct into the *psf1-1* strain and its wild type parent strain and assayed splicing by primer extension and copper growth (Figure 2B, Table 1). If the *psf1-1* mutation only affects polarity, then splice site usage in the +A WT construct would have shifted towards the distal 3' splice site as it does in the +T PyDOWN construct. Instead, we found that 3' splice site usage was shifted away from the uridine-rich, distal 3' splice site and towards the uridine-poor, proximal splice site with +A WT (wild type proximal/distal = .86; *psf1-1* proximal/distal = 3.1). The *psf1-1* mutant is, therefore, impaired in uridine recognition at the 3' splice site and not in "polarity" of splice site choice.

To confirm these results and rule out context specific effects, we tested an isogenic set of constructs, +TA and +AT (Figure 1A). These constructs are similar to +T PyDown and +A WT except that the adenosine-rich and uridine-rich tracts preceding the competing 3' splice sites are uniform and identical. Primer extension analysis of +TA and +AT introduced into wild type and *psf1-1* strains yields the same qualitative results as seen with +T PyDown and +A WT (Figure 2C and 2D, Table 1), i.e. the uridine-rich 3' splice site is less favored in the *psf1-1* mutant than in the wild type strain. The total amount of mature message produced with these constructs is markedly lower in the *psf1-1* mutant versus wild type. Importantly, however, the influence that the *psf1-1* mutation exerts on utilization of uridine-rich 3' splice sites is very similar for both sets of constructs (Table 1, Loss of uridine preference for +T PyDOWN and +TA; +A WT and +AT). Thus, the *psf1-1* mutant is specifically impaired in recognizing the uridine-rich tract at the 3' splice site.

Splicing of the TUB3 *intron requires Psf1*

The altered pattern of splice site choice in the *psf1-1* strain prompted us to determine whether splicing of endogenous yeast genes is affected by this mutation. While most yeast 3' splice sites are enriched for uridines (Parker & Patterson, 1987; Rymond & Rosbash, 1992), mutations in the uridine tract often have no apparent phenotype suggesting that recognition of this sequence is not usually rate-limiting for 3' splice site selection (Patterson, 1989). Primer extension analysis of several endogenous spliced yeast genes (*RP51, CYH2, UBI1*) revealed no apparent splicing defects in the *psf1-1* strain (data not shown). We then examined the splicing of an unusual intron in the a tubulin homologue, *TUB3*. The distance between the branchsite and 3' splice site of this intron is atypically long (139 nucleotides) for yeast and the 3' splice site is highly enriched for uridines (Figure 3A) (Schatz et al., 1986a). A non-uridine-rich spacer sequence of similar length, when introduced into the 3' splice site of the yeast actin intron, strongly inhibits its splicing (Cellini et al., 1986). We reasoned that the *TUB3* intron, because of its unusual architecture, might be highly dependent on its uridine tract, and thus the *PSF1* gene product, for efficient splicing. Since *TUB3* is expressed at relatively low levels, we chose to move the intron into the well expressed *ACT1-CUP1* gene fusion construct to facilitate analysis. As shown in figure 3A, this construct (TUB3 WT) contains the actin 5' exon and 5' splice site, most of the *TUB3* intron (except for the 5' splice site region), and the 3' exon encoding *CUP1*. For comparison, splicing of the wild type *ACT1-CUP1* gene fusion is shown in lanes 1 and 2 of Figure 3B. The *TUB3* intron is spliced as efficiently as the wild type *ACT1-CUP1* fusion in a wild type strain, whereas it is spliced poorly in the *psf1-1* mutant as evidenced by primer extension (mature message/lariat intermediate diminished 20 fold), and by copper growth (copper resistance diminished 8 fold in *psf1-1* mutant)(Figure 3B, lanes 3 and 4; Table 2). The block to the second step of splicing indicated by the decreased ratio of (mature message/lariat intermediate) is expected for a mutant impaired in 3' splice site recognition. We observe no change in TUB3 WT pre-mRNA levels indicating that the splicing defect in the *psf1-1* strain occurs exclusively at the second catalytic step (data not shown). The splicing of the actin intron, which has a more typical spacing of 43 nucleotides between its branchsite and 3' splice site, is only mildly affected (mature message/lariat intermediate diminished 3 fold) in the *psf1-1* strain (Figure 3B, lanes 1 and 2; Table 2).

To establish that impaired recognition of the uridine-rich tract, and not another feature of the *TUB3* intron, is the basis for the *psf1-1* splicing defect, we mutated the 3' splice site uridine tract in our *TUB3* fusion construct to adenosines (TUB3 polyA) (Figure 3A). We predicted that without a uridine-rich 3' splice site, the splicing of the TUB3 polyA intron would be impaired, and would be insensitive to the presence of the *psf1-1* mutation. This construct was introduced into wild type or *psf1-1* mutant strains and analyzed by primer extension and copper growth. Indeed, as predicted, splicing of the TUB3 polyA intron is almost completely insensitive to the presence of the *psf1-1* mutation (Figure 3B, lanes 5 and 6; Table 2). Furthermore, in a wild type strain, the TUB3 polyA intron is spliced more poorly than is the TUB3 WT intron (Figure 3B, lane 3 versus 5; Table 2). Surprisingly, however, the TUB3 polyA intron is spliced better in a *psf1-1* strain than is the TUB3 WT intron (Figure 3B, lane 4 versus 6). Therefore, it appears that adenosines, though not preferred, can be recognized by the splicing machinery when uridines are not present at the 3' splice site. Moreover, the ability to utilize adenosine-rich 3' splice sites must be independent of the uridine recognition mediated by Psf1, since the TUB3 polyA intron is spliced at similar efficiencies in both wild type and *psf1-1* strains. Thus, despite our finding that adenosines at the 3' splice site can be recognized better than expected, this experiment demonstrates that loss of uridine-tract recognition is the basis of the *TUB3* splicing defect in the *psf1-1* strain.

Additionally, we found that *psf1-1* mutants are hypersensitive to the microtubule destabilizing drug benomyl at 15 μg/ml (data not shown). A *TUB3* gene disruption also causes benomyl hypersensitivity (Schatz et al., 1986b), indicating that our observations of the *TUB3* fusion construct very likely reflect decreased expression of the endogenous *TUB3* gene. *psf1-1* mutants also display mild temperature sensitive and cold sensitive growth phenotypes (data not shown); *TUB3* null mutants do not display these additional phenotypes, suggesting that besides *TUB3*, there are other highly "uridine-dependent" introns in which splicing is affected by the *psf1-1* mutation.

psf1-1 *Exacerbates the Defects of 3' Splice Site PyAG Mutations*

Although uridine recognition by Psf1 clearly plays a role in splice site selection and in the efficient splicing of introns such as that in the *TUB3* gene, it might also make a contribution to the splicing of other introns. This contribution would normally be masked if 3' splice site recognition

were not the rate-limiting step in splicing these introns. Therefore, to test for a more general role of uridine recognition by Psf1 in 3' splice site selection, we constructed point mutations in the PyAG motif of the actin 3' splice site. Point mutations in this sequence partially or completely block the second catalytic step of splicing and make 3' splice site utilization rate-limiting for production of mature message (Vijayraghavan et al., 1986; Fouser & Friesen, 1987). In this situation, one could imagine that recognition of the uridine-rich tract might be critical, even for an apparently "uridine-independent" intron. Thus, a defect in the *psf1-1* strain would be apparent in the splicing of these 3' splice site mutant RNAs.

We tested a number of actin 3' splice site mutations in the *ACT1-CUP1* reporter. Splicing of three representative mutants (TAG to GAG, TTG or TGG) and the wild type intron was analyzed by primer extension and copper growth (Figure 4, Table 3). As shown in lanes 1 and 2 of Figure 4, the splicing of the wild type actin intron is only mildly affected by the *psf1-1* mutation. As expected, in a wild type strain, all three intron mutations result in increased lariat intermediate and reduced amounts of mature message, indicating their partial defect in the second catalytic step of splicing. Moreover, in a *psf1-1* strain the mutant introns show an even greater decrease in splicing efficiency (5 to 10 fold reduction in mature message/lariat intermediate). In addition, we observe a similar effect when substitutions in the uridine tract are made in cis with PyAG mutations (data not shown). Similar to*TUB3* pre-mRNA splicing, we see no change in pre-mRNA levels for these constructs in the *psf1-1* strain, indicating that the impairment of splicing is at the second catalytic step (data not shown). In summary, Psf1-dependent uridine tract recognition can make a significant contribution to splicing efficiency, even in the actin intron which does not normally depend on its uridine tract for efficient splicing.

PSF1 *is identical to* PRP8

In order to determine the molecular basis for the *psf1-1* phenotype, we isolated genomic clones of *PSF1* and *psf1-1*. Wild type *PSF1* clones were obtained from a standard yeast plasmid library which was transformed into a *psf1-1* strain harboring the TUB3 WT construct. A wild type copy of the *PSF1* gene was expected to improve the splicing efficiency of the TUB3 WT RNA in a *psf1-1* strain, and thus confer increased copper resistance (see Methods).

Because the *psf1-1* mutant causes a partial splicing phenotype in the presence of a wild type *PSF1* allele (Table 1, data not shown), to isolate the *psf1-1* allele, we constructed a plasmid library from the mutant strain and used it to transform a wild type strain harboring the +T PyDOWN reporter construct (see Methods). A *psf1-1* clone introduced into this strain is predicted to activate use of the distal splice site and, therefore, confer increased copper resistance.

Both strategies yielded clones with overlapping restriction maps. Partial sequence analysis revealed that the cloned inserts encoded the U5 snRNP protein Prp8 (Lossky et al., 1987; Jackson et al., 1988). Both clones complemented a chromosomal deletion of the *PRP8* gene; however, only the *psf1-1* clone could activate the distal splice site in the +T PyDOWN construct, as expected (data not shown). The *psf1-1* mutation was mapped using gap repair (Guthrie & Fink, 1991) and shown to reside in the carboxyl-terminal portion of the gene. Sequence analysis of this region revealed a single G to A transition in the *psf1-1* clone that changes a GAG glutamic acid codon to an AAG lysine codon at amino acid position 1960. Subsequent analysis of *psf1-2* revealed the identical nucleotide change.

To demonstrate that this G to A mutation causes the *psf1-1* phenotype, a small, sequenced restriction fragment containing the mutation was used to replace the corresponding region of the wild type gene (see Methods). The swap produced a single point mutation in an otherwise wild type *PRP8* clone and was both necessary and sufficient to generate the *psf1-1* phenotypes. The formal designation of *psf1-1* will be *prp8-101*, but for clarity, we will continue to refer to it as *psf1-1* in this report.

Allele Specificity of the psf1-1 *Phenotype*

The identity of *PSF1* with *PRP8* raised the question of whether the *psf1-1* phenotype is unique to this allele or commonly associated with *prp8* mutants. The low frequency of *psf* mutants in our genetic selection and the fact that both independently isolated *psf1* alleles contain identical mutations suggest that *psf1-1* confers a novel phenotype. To test this directly, we introduced the +T PyDOWN reporter construct (Figure 1) into *prp8* strains carrying either one of two temperature sensitive alleles, *prp8-1* and *prp8-7*, or a cold sensitive allele, $prp8^{cs}$ (see Methods for strain references). RNA was prepared from these strains grown at a permissive temperature (25° C) and the splicing products were analyzed by primer extension (Figure 5). Use of the alternative 3'

splice sites in the +T PyDOWN construct was essentially the same as in the wild type strain for these alleles. Only the *psf1-1* allele of *PRP8* causes loss of uridine recognition and consequent activation of the distal splice site with this construct. Thus, the *psf1-1* mutation defines a novel function for the Prp8 protein.

Site-Specific UV Crosslinking of Prp8 to the 3' Splice Site

The simplest model to explain the role of Prp8 in 3' splice site selection is that Prp8 directly binds and recognizes the 3' splice site. To test this model, we used the technique of Moore and Sharp (1992) to introduce a single 32-phosphate label at the 3' splice site of an *in vitro* splicing substrate (XL7) for the purpose of site-specific UV crosslinking (Figure 6A). XL7 is a derivative of the *ACT1-CUP1* gene fusion with the 5' exon and intron containing 50% uniformly incorporated 5-bromo-uridine as a photoaffinity reagent (Hanna, 1989). We also introduced a A to G point mutation (TAG to TGG) in the penultimate residue of the intron (Figure 6A). The effect of this mutation, *in vivo* as well as *in vitro*, is to slow the normally rapid kinetics of the second catalytic step of splicing without changing 3' cleavage site choice.

For *in vitro* splicing, we employed a novel method for preparing splicing extracts. The procedure, which involves grinding cells in liquid nitrogen, allows rapid and reproducible preparation of highly active whole cell splicing extracts (see Methods). Extracts were prepared from two isogenic *prp8* chromosomal deletion strains. One harbors a wild type *PRP8*-containing plasmid and the other harbors a plasmid with an epitope tagged version of *PRP8*, containing three copies of a hemagglutinin epitope at its carboxyl terminus (Guthrie & Fink, 1991)(Methods). The epitope tagged version of *PRP8*, *PRP8-HA3*, complements a chromosomal deletion of *PRP8* and supports wild type growth rates at all temperatures tested (16°-37° C)(data not shown).

XL7 undergoes the first catalytic step of splicing efficiently, accumulates lariat intermediate and shows a reduced amount of mature message, indicated by the excised lariat intron (Figure 6B, lanes 1 and 3). The placement of the labeled phosphate in the intron (Figure 6A) precludes visualization of mature spliced RNA and the free first exon. As expected, XL7 was not spliced in the absence of ATP (Figure 6B, lane 2).

A portion of each sample shown in figure 6B was subjected to UV crosslinking followed by digestion with RNase T1. The samples were then analyzed by autoradiography following SDS-

PAGE, before or after immunoprecipitation using a monoclonal antibody specific for the hemagglutinin epitope. Numerous polypeptides became crosslinked to the XL7 3' splice site in splicing extract, most of them in an ATP independent manner (Figure 6C, lanes 1 and 3 versus lane 2). The crosslinks were all dependent on irradiation with UV light (data not shown). Immunoprecipitation of Prp8-HA3 revealed a single labeled band that migrated at the molecular weight of Prp8 (280 kd)(Figure 6C, lane 4). Furthermore, appearance of this band was dependent on ATP and the presence of an epitope tag on Prp8. These results demonstrate that the crosslinked species corresponds to Prp8 (Figure 6C, lane 4 versus 6) and that Prp8 associates with the 3' splice site in a splicing dependent manner (Figure 6C, lane 4 versus 5).

Because Prp8 may make multiple contacts with the splicing substrate, incomplete RNase T1 digestion combined with crosslinking at a more distant site in the substrate would give the appearance of specific crosslinks in our experiments. To confirm that the RNase T1 digestion we performed was complete, the immunoprecipitated material from a sample identical to that in Figure 6C (lane 4) was digested with protease and extracted (see Methods). The RNA was precipitated from this sample and size fractionated (Figure 6D). The labeled RNA in Figure 6D appears slightly larger and more heterogeneous than the 15 nucleotide fragment from digestion of pure XL7 RNA with RNase T1 (data not shown). This size distribution, slightly above the 15 nucleotide size marker, is probably due to peptide adducts which remain covalently attached to the RNA after digestion and extraction. Importantly, no RNAs larger than these are present in the immunoprecipitate indicating that the 3' splice site is indeed the site of crosslinking detected in our experiments.

To demonstrate further that the interaction between Prp8 and the 3' splice site is splicing specific, we tested crosslinking using derivatives of XL7 that contain additional mutations. XL7-A5 contains a G to A mutation in the fifth position of the intron. This mutation strongly blocks the first step of splicing (Figure 6E, lane2) and results in no detectable Prp8-3' splice site crosslinking (Figure 6F, lane 2). XL7-C259 contains an A to C mutation at the branchsite A residue. This substrate is spliced five to ten fold more poorly than XL7 (Figure 6E, lanes 1 and 3) and shows a similar reduction in Prp8-3' splice site crosslinking (Figure 6F, lanes 1 and 3). The strong correlation between the levels of splicing and crosslinking establishes that Prp8 interacts with the 3' splice site in active spliceosomes.

To address the possibility that Prp8 interacts non-specifically with all regions of the substrate during splicing, we examined crosslinking in a derivative of XL7, XL7-E2, that contains a uniformly labeled second exon (The first labeled RNase T1 fragment begins 9 nucleotides downstream of the 3' splice site; see Methods). With this substrate we see no detectable crosslinks to Prp8 (Figures 6E and F, lanes 4 and 5). When the specific activities of individual RNase T1 fragments are taken into account, the assay could detect crosslinking to any fragment in exon two, even with a ten to fifty fold reduction in efficiency compared with crosslinking to XL7. Therefore, Prp8 does not interact promiscuously with all regions of the substrate during splicing; it interacts specifically with the 3' splice site (and perhaps other conserved sequences; see Discussion).

Discussion

Prp8 Promotes use of the Uridine-Rich Tract at the 3' Splice Site

The conserved uridine-rich tract upstream of the 3' splice site in yeast is important for efficient utilization of that splice site (Patterson & Guthrie, 1991). Here, we have presented evidence for a functional interaction between Prp8 and the uridine-rich tract. Using a sensitive cis-competition assay, we have identified the *psf1-1* mutant allele of *PRP8* based on its loss of preference for a uridine-rich 3' splice site and consequent activation of a non-preferred, competing, adenosine-rich 3' splice site. This phenotype contrasts that of the *slu7-1* mutant which is impaired in utilization of branchsite-distal 3' splice sites independent of sequence content (Frank & Guthrie, 1992). *slu7-1*, therefore, appears to affect spacing constraints or "polarity". The relationship between Prp8 and Slu7 is discussed below.

Because yeast introns do not normally contain competing 3' splice sites, we searched for endogenous yeast introns that might require the uridine recognition function of Prp8 and thus be affected by the *psf1-1* mutation. The *TUB3* intron is a good candidate for such an intron because of its unusual architecture; it has an atypically long branchsite to 3' splice site distance (139 nts.) and a high enrichment for uridines preceding its 3' splice site (10/15 residues preceding the PyAG are uridines). Indeed, the *TUB3* intron is spliced very poorly in the *psf1-1* strain. Furthermore, the effect of *psf1-1* on the splicing of the *TUB3* intron can be obviated when the uridine tract is

mutated to adenosines. Thus, the splicing defect of the *TUB3* intron in a *psf1-1* strain can be attributed to loss of uridine recognition.

To determine if the uridine recognition function of Prp8 contributes to the splicing of other types of introns, albeit to a lesser extent, we made 3' splice site selection rate-limiting in the actin intron by introducing point mutations in the PyAG motif. In this situation the *psf1-1* mutation exerts a large deleterious effect on splicing, demonstrating the contribution of uridine tract recognition by Prp8, even in the actin intron, whose architecture is typical of yeast introns. Thus, uridine tract recognition by Prp8 probably contributes to the splicing efficiency of many yeast introns.

The psf1-1 *Mutation Defines a Novel Function for Prp8*

Previous studies of Prp8 have implicated this protein in early steps of spliceosome assembly. Depletion of Prp8 or inactivation of a temperature sensitive mutant, *prp8-1*, leads to an early block in spliceosome formation, at a step prior to U4, U5 and U6 triple snRNP addition. However, the function(s) of Prp8 in the spliceosome, subsequent to the spliceosome assembly block, could not be assessed by these methods (Brown & Beggs, 1992).

Our genetic approach has allowed us to identify a novel function for Prp8 by demanding a precise functional defect, loss of preference for a uridine-rich 3' splice site. The *psf1-1* phenotype most likely results from a specific inactivation of the domain(s) in Prp8 responsible for uridine-rich 3' splice site recognition rather than a general inactivation of the protein. Supporting this is our finding that the uridine recognition phenotype is specific for the *psf1-1* allele of *PRP8*. Most notably, unlike other alleles of *PRP8* that are blocked prior to the first catalytic step of splicing, the biochemical block caused by the *psf1-1* mutation is at the second catalytic step of splicing, which is when 3' splice site selection takes place. Moreover, other *prp* mutants that are specifically defective for the second step of splicing do not display the uridine recognition defect of *psf1-1* (Frank & Guthrie, 1992). Finally, the novel phenotype of the *psf1-1* allele provides an important tool for the dissection of the 280 kD Prp8 protein into structure:function modules. As a first step, we are isolating additional alleles of *PRP8* which confer a *psf1*-like phenotype.

Prp8 Interacts Directly with the 3' Splice Site

The above observations of the *psf1-1* mutant phenotype are consistent with a mechanism whereby Prp8 binds directly to the uridine tract of a 3' splice site, activating its use. To test this model, we used site-specific UV crosslinking to determine whether the Prp8 protein contacts the 3' splice site during the splicing reaction. Because the kinetics of the second catalytic step of splicing are rapid (Schwer & Guthrie, 1991), we wished to slow down this reaction in order to accumulate possible intermediates in 3' splice site recognition. The substrate we used contains an A to G point mutation in the dinucleotide splice acceptor (AG to GG), and undergoes the first catalytic step of splicing efficiently, accumulating lariat intermediates. The second step of splicing occurs accurately but at a reduced rate. With this substrate, we were able to identify a UV induced crosslink between Prp8 and a 15 nucleotide region of the intron that contains the uridine-rich tract and the 3' splice site. This crosslink is splicing specific as it is absent when ATP is omitted and correlates with the reduced levels of splicing activity when mutant substrates are utilized. The absence of crosslinking between Prp8 and the uniformly labeled second exon indicates that the interaction between Prp8 and the 3' splice site is specific. The demonstration of a direct physical contact between Prp8 and the 3' splice site, combined with the genetic data above, provides strong support for a model in which Prp8 binds and activates uridine containing 3' splice sites (see below and Figure 7).

Our results are interesting in light of previous work that demonstrated crosslinking between Prp8, or its mammalian homologue, p220, and the splicing substrate (Garcia-Blanco et al. 1990; Whittaker and Beggs 1991). Although these crosslinks were not mapped, subsequent work identified the -2 position of the first exon (relative to the 5' splice site) as one p220 crosslinking site (Wyatt et al., 1992). The functions of these Prp8/p220-pre-mRNA interactions have not yet been determined, but could reflect an earlier event involving the U5 snRNP (see below). Given the conservation of the polypyrimidine tract and, in particular, the sequence and large size of the *PRP8* gene (Anderson et al. 1989; Pinto and Steitz 1989; Garcia-Blanco et al. 1990; J.G.U. and C.G. unpublished observations), it is likely that the 3' splice site recognition function for Prp8 that we have identified in yeast is similar in mammals and other organisms. Consistent with this, experiments in mammalian extracts suggest a role for the polypyrimidine tract late in the splicing pathway, similar to one that has been found in yeast (Reed, 1989). As in yeast, the

polypyrimidine tract may not be critical for the second step of splicing in all mammalian introns (Smith et al., 1989; Smith et al., 1993).

The Role of U5 snRNP in 3' Splice Site Selection and the Second Catalytic Step of Splicing

In yeast, three trans-acting factors have now been shown to affect 3' splice site selection. They are Prp8, which interacts with the conserved uridine-rich tract at the 3' splice site (this study); the conserved loop in U5 snRNA, which can interact with non-conserved sequences in the second exon adjacent to the 3' splice site (Newman & Norman, 1992; Sontheimer & Steitz, 1993); and Slu7, which promotes use of branchsite-distal 3' splice sites in a sequence non-specific manner (Frank & Guthrie, 1992). The former two are components of U5 snRNP, while Slu7 interacts genetically with U5 snRNA.

How might these factors function to promote 3' splice site recognition and the second catalytic step of splicing? The conserved loop of U5 snRNA and Prp8 are both in the same snRNP particle and are, therefore, very likely to interact with the 3' splice site region simultaneously. Although neither of these interactions would normally provide the necessary specificity for 3' splice site selection, together, they might serve to enhance recognition of a properly situated AG dinucleotide which is bound by an, as yet, unidentified factor. Our observation of a synergistic inhibition of splicing of 3' splice site PyAG mutants in a *psf1-1* strain is consistent with the idea that Prp8 and this factor bind the 3' splice site cooperatively. How Slu7 functions in this process is less clear. Based on the phenotype of *slu7-1*, Frank and Guthrie (1992) have proposed that Slu7 interacts with the splicing factor that directly recognizes the uridine-rich tract at the 3' splice site. If so, then Slu7 might serve to monitor or promote the binding of Prp8 to the 3' splice site.

As described above, the mammalian Prp8 homologue, p220, and the conserved loop of U5 snRNA have also been shown to interact with the first exon during splicing (Newman & Norman, 1991; Newman & Norman, 1992; Wyatt et al., 1992; Cortes et al., 1993; Sontheimer & Steitz, 1993). These findings have implicated Prp8 and U5 snRNP as playing an important role during the first catalytic step. Taken together with our results, Prp8 thus appears to promote both catalytic steps of the splicing reaction. Furthermore, if the interactions between U5 snRNP and exon one are maintained after the first catalytic step, then Prp8 and U5 snRNA would each bind the 5' exon and 3' splice site simultaneously; in this case, Prp8 could serve to juxtapose the reaction partners for the second catalytic

Materials

The following materials were purchased from the indicated sources. Restriction enzymes, New England Biolabs; T4 polynucleotide kinase, Pharmacia; T4 DNA ligase and Sequenase, U. S. Biochemical; MLV Reverse Transcriptase, Bethesda Research Laboratories; 5FOA, PCR Incorporated; EMS and 5-bromo-UTP, Sigma; [g-32P]ATP, [a-32P]ATP and Hot Tub DNA polymerase, Amersham; 12CA5 antibodies, Babco; RNase T1, Boehringer; Oligonucleotides were synthesized by Hiten Madhani, Suzanne Noble, Yan Wang and Chris Siebel on a MilliGen Plus Cyclone DNA Synthesizer. Benomyl was a generous gift from Du Pont Chemicals.

Genetic Methods

All genetic manipulations including EMS mutagenesis were carried out as described in (Guthrie & Fink, 1991). For EMS mutagenesis, ~10^9 haploid cells from strain L5 containing *ACT1-CUP1* +T PyDOWN were mutagenized to 50% lethality and plated directly on media containing .18 or .25 mM $CuSO_4$. After four days growth at 30°, colonies were picked and mated to strain I4. Dominant mutants were discarded and the remaining strains were cured of the reporter plasmid and retested for copper growth. Those that did not behave as *cup1* null mutants (.013 mM maximum copper resistance) without the reporter plasmid were also discarded. Two mutants strains fulfilled these criteria. A mating test and subsequent analysis (see Results) showed that the two mutants, *psf1-1* and *psf1-2*, were allelic. Copper growth assays were performed as described previously (Lesser & Guthrie, 1993).

Strains

The following *S. cerevisiae* strains were used in this study:

L5: *MATa cup1Δ::ura3 leu2 trp1 ura3-52 lys2 his3 ade* (derivative of K3 [Lesser & Guthrie, 1993] constructed by J. Lefstin).

I4: *MATa cup1Δ::URA3 leu2 trp1 ura3-52 lys2 his3 ade* (Lesser & Guthrie, 1993).

YJU69: Identical to L5 except it contains a chromosomal *psf1-1* (*prp8-101*) mutation.

N13/L5: Diploid formed from mating of YJU69 and the MATα version of L5.

JDY8.02: *prp8Δ::LEU2 MATa leu2-3 leu2-112 tyr1 his3Δ1 ura3-52 ade2-101 trp1-289* pY8000 (*URA3* marked *PRP8* plasmid) (Brown & Beggs, 1992).

YJU74: *prp8Δ::LEU2 MATa leu2-3 leu2-112 tyr1 his3Δ1 ura3-52 ade2-101 trp1-289 pep4Δ::TRP1* pY8000

YJU76: Identical to YJU74 except contains *PRP8* plasmid pJU186.

YJU77: Identical to YJU74 except contains *PRP8-HA3* plasmid pJU204.

rna8-1: *prp8-1 MATa ade arg trp1 ura3* (Gift from Anita Hopper, Penn. State Univ. School of Medicine).

DJY76: *prp8-7 MATa leu2-3 leu2-112 tyr1 his3 ura3-52* (Gift from Jean Beggs, Univ. of Edinburgh).

YSN157: *prp8cs MATa ura3 leu2 his3 lys2 ade 2* (Gift from Suzanne Noble, UCSF).

E. coli strain DH5α was used as a recipient for all cloning procedures.

Plasmids

Standard cloning procedures were used for all plasmid constructions (Maniatis et al., 1982). pGAC14 (*ACT1-CUP1*)(Lesser & Guthrie, 1993) was the starting plasmid for all reporter constructs used here. Derivatives were made by PCR mutagenesis. TUB3 WT (*ACT1-TUB3-CUP1*) was made by replacing the pGAC14 Xho1-Kpn1 fragment with *TUB3* intron nucleotides 67-298 amplified from plasmid pRB300 (Schatz et al., 1986a). Other *ACT1-TUB3-CUP1* and *ACT1-CUP1* derivatives were

constructed by replacing the BamHI-KpnI fragment from one of these constructs with mutant PCR fragments.

PRP8 plasmids were constructed from the YCP50 *PRP8* or *psf1-1* clones (see below) by cutting with XhoI and NheI and ligating the gene into the XbaI and SalI sites of the HIS3-CEN-ARS vector pSE362 (Elledge & Davis, 1988) to generate pJU186 and pJU187. PCR mutagenesis was used to introduce a unique NotI restriction site into the C-terminus of pJU186 just before the termination codon. A small restriction fragment containing three tandem copies of the hemagglutinin epitope (Tyers et al., 1992) was inserted into this site to generate pJU204 *(PRP8-HA3).*

Cloning psf1-1 *and* PRP8

The wild type *PSF1 (PRP8)* gene was cloned by complementing the copper growth defect of a *psf1-1* strain containing the TUB3 WT reporter using a YCP50 based genomic library (gift from Mark Rose, Princeton Univ.). Transformants were selected for growth at 1 mM copper. The mutant *psf1-1 (prp8-101)* allele was cloned by constructing a library of Sau3a partially digested *psf1-1* genomic DNA which was inserted into the BamHI site of YCP50. This library was transformed into strain L5 which contained the +T PyDOWN reporter. Transformants were screened for growth at .1 mM copper. Three wild type and twelve *psf1-1* mutant clones containing overlapping inserts were recovered (Dower et al., 1988; Strathern & Higgins, 1991). A mutant clone was gap repaired to map the *psf1-1* mutation and a 1.9 Kb ClaI fragment containing the mutation was sequenced to find the change. The single point mutation was re-introduced into the wild type clone pJU186 by replacing its 1.9 Kb ClaI fragment with the mutant version.

RNA Preparation and Primer Extension Analysis

Cells were grown to mid-logarithmic phase to harvest RNA. Total RNA was prepared according to Lesser and Guthrie (1993). The exon 2 and U1 primers used were: Cup 24mer 5'-CTTCATTTTGGAAGTTAATTAATT-3'; and 19K, 5'-CAATGACTTCAATGAACAATTAT-3'. Primer extensions were performed by annealing 1 ng of 5' end labeled primer to 15 µgm total RNA in TE with .25M KCl in 10 µl. Annealing was performed by slow cooling the mixture from 70° to 42° at which point 25 µl RT mix (15 mM Tris-HCl pH 8.3, 7.5mM MgCl2, 75 µgm/ml actinomycin D, .625 mM dNTPs, 1 mM DTT, 4U/µl MLV reverse transcriptase) was added. Samples were incubated at

42° for 30 minutes and ethanol precipitated. cDNAs were fractionated on 6% acrylamide/7M urea/TBE sequencing gels.

RNA substrates

Templates for *in vitro* transcription were generated by PCR using the PG14 *ACT1-CUP1* plasmid as a target for amplification. Oligonucleotide primers for the 5' half molecule were T7-5'II, 5'-GCGCTAATACGACTCACTAT-AGGGAGATTTTTCACGCTTAC-3' and XL7, 5'-GCGGTACCAAGGAAAAATATAGCA-ACAAAAAGAATGAAG-3'. Oligonucleotides for the 3' half molecule were XL1, 5'-GCGCTAATACGACTCACTATAGGGTTGCTGGTACCGGGT-3' and CUPE2HindIII, 5'-GGGGAAGCTTACATTGGCATTGGCACTC-3'. The 5' half template was cut with StyI and the 3' half template was cut with HindIII prior to transcription. RNAs were synthesized using an Ambion T7 Megascript kit according to instructions. For the 5' half RNA, 50% of the UTP was substituted with 5-bromo-UTP. RNAs were gel purified, end labeled and ligated as described previously (Wyatt et al., 1992), with each RNA at 2 μM in a 10 or 20 μl ligation reaction. The bridging deoxyoligonucleotide, 5'-ACCCGGTACCAACCCCAAGGAAAAATATAGCAA-3', was also at 2 μM. Efficiency of ligation reactions was 10%-20%. XL7-A5 and XL7-C259 were constructed as above except that the appropriate mutant PCR templates were utilized. XL7-E2 was constructed by labeling the 3' half RNA uniformly with a32P-ATP and incorporating 50% 5-bromo-UTP. The first labeled RNase T1 fragment begins nine nucleotides downstream of the 3' splice site and the fragments range in size between 2 and 21 nucleotides. The 21 nucleotide fragment contains half of the 32P label.

Splicing Extract Preparation

Whole cell extracts were prepared using an adaptation of a previously described method (Sorger et al 1989). 1-3 liters of cells from strains YJU76 and YJU77 were grown to late logarithmic phase in YEPD. Cells were harvested by centrifugation and washed twice with 50 and 20 mls buffer AGK (10 mM HEPES pH 7.9, 1.5 mM MgCl2, 200 mM KCl, 10% vol/vol glycerol). Cell pellets were resuspended in .4 cell pellet volumes AGK with protease inhibitors (final concentrations 1 mM PMSF, 1 μg/ml leupeptin, 1 μg/ml aprotinin, 1 mM benzamidine). The suspension was added to a 5" porcelain mortar (Fisher) containing liquid nitrogen and ground to a fine powder with occasional addition of liquid nitrogen. The frozen powder was transferred to a centrifuge tube, thawed rapidly at

25° or 30°, and centrifuged in a SS34 rotor at 17K, 4° for thirty minutes. From this point, the samples were treated as described previously for yeast splicing extract preparation except that the dialysis buffer volume was 1.5 or 2 liters and was changed once during the procedure (Lin et al., 1985).

in vitro *Splicing and UV Crosslinking*

Standard *in vitro* splicing reactions (Lin et al., 1985) were in 40 µl with $3x10^5$ cpm substrate. After 30 minutes incubation at 25°, a small aliquot was removed for analysis of the RNA. The remainder was placed in a microtiter dish on ice and irradiated at 312 nm with a hand held UV lamp (Fisher) for 5 minutes at a distance of .5-1 cm. Samples were removed and protease inhibitors (see above) and RNase T1 (10 units/µl final) were added and the samples incubated at 37° for 30 minutes. An aliquot of 5 µl was removed and added to SDS-PAGE sample buffer for total protein analysis. KCl and NP40 were added to the remaining sample to final concentrations of 300 mM and .05% respectively, and 1 µg protein A purified 12CA5 antibody was added to each sample. After 1 hour on ice, protein A sepharose beads (Pharmacia) were added in 500 µl of KTN buffer (.3 M KCl, 20 mM Tris pH 7.4, .05% NP40) and the samples gently rocked at 4° for an additional hour. The protein A sepharose beads were washed three times with KTN buffer and sample buffer was added to the pellet. After boiling for two minutes, the samples were fractionated on a 7.5% SDS-polyacrylamide gel and autoradiographed. For analysis of the crosslinked RNA fragment, the pellet from the immunoprecipitate was resuspended in 40 µl 1% SDS, .5 mg/ml proteinase K (Sigma), 10 mM EDTA and incubated at 50° for thirty minutes. The sample was then extracted twice with phenol/chloroform (1:1), ethanol precipitated and fractionated on a 20% acrylamide/TBE/7M urea gel and autoradiographed.

Acknowledgments

We thank Jeremy Brown, Shelly Jones, Hiten Madhani, Maho Niwa, Suzanne Noble, Chris Siebel, Jon Staley and Peter Walter for comments on this manuscript. We thank Lucita Esperas and Carol Pudlow for their excellent technical assistance. We thank Jean Beggs for providing PRP8 sequence information prior to its submission to Genbank.

This work was supported by a NIH training grant, a NSF pre-doctoral fellowship to J.G.U., and NIH research grant GM21119 to C.G. C.G. is an American Cancer Society Research Professor of Molecular Genetics.

Figure 1. Splice Site Competitions and Rationale for Isolating *psf* Mutants.

(A) Sequence of 3' splice site region of competition constructs used in this study. The branch nucleotide is raised and underlined and the 3' splice acceptors are underlined.

(B) Rationale for use of +T PyDOWN competition in selecting *psf* mutants. The top figure diagrams hypothetical PSF protein binding and activating the uridine-rich proximal splice site in +T PyDOWN RNA. The bottom figure shows reduced binding of the mutant psf protein and activation of splicing to the distal 3' splice site. Arrows indicate relative use of the two competing splice sites in the wild type and mutant strains.

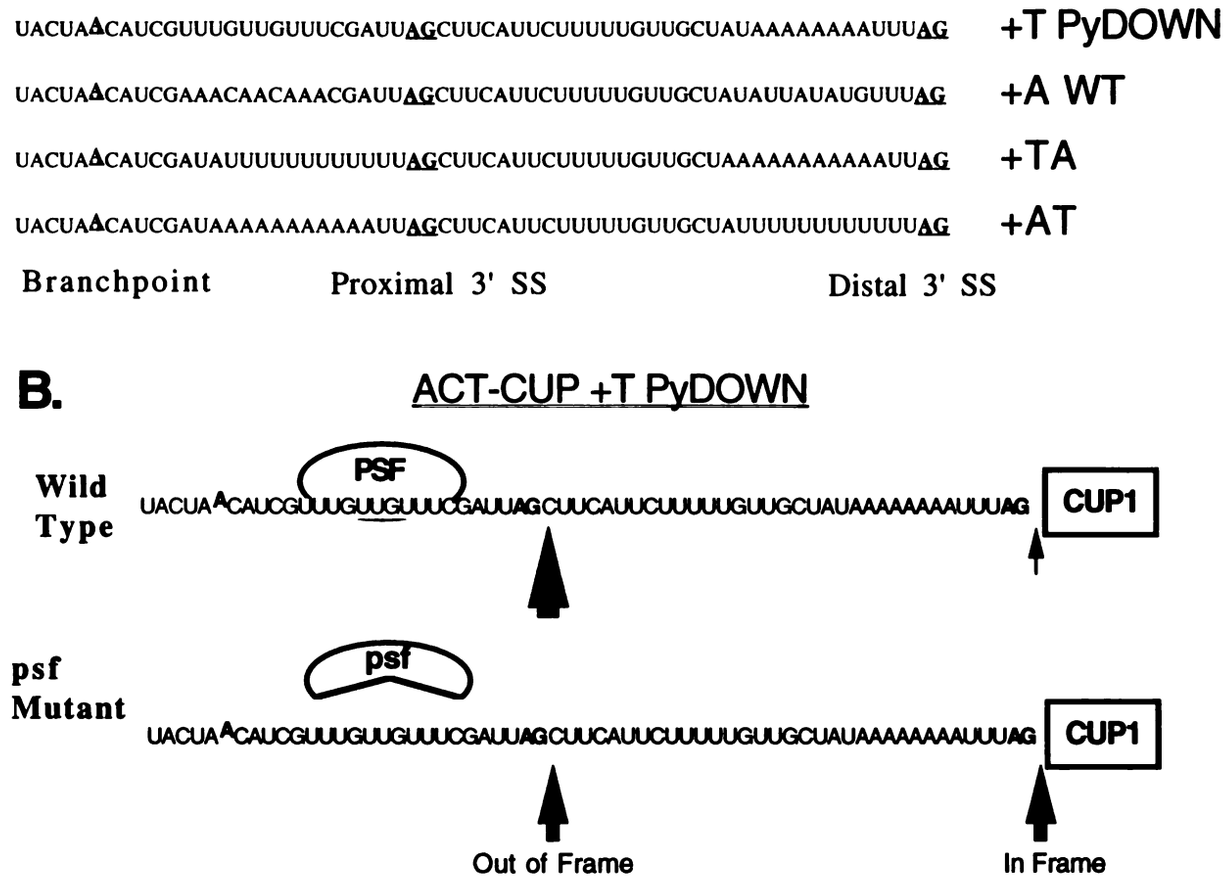

A.
UACUAACAUCGUUUGUUGUUUCGAUUAGCUUCAUUCUUUUUGUUGCUAUAAAAAAAAUUUAG
+T PyDOWN
UACUAACAUCGAAACAACAAACGAUUAGCUUCAUUCUUUUUGUUGCUAUAUUAUAUGUUUAG
+A WT
UACUAACAUCGAUAUUUUUUUUUUUUAGCUUCAUUCUUUUUGUUGCUAAAAAAAAAAAAUUAG
+TA
UACUAACAUCGAUAAAAAAAAAAAAUUAGCUUCAUUCUUUUUGUUGCUAUUUUUUUUUUUUUAG
+AT
Branchpoint
Proximal 3' SS
Distal 3' SS
B.
ACT-CUP +T PyDOWN
Wild Type
PSF
UACUAACAUCGUUUGUUGUUUCGAUUAGCUUCAUUCUUUUUGUUGCUAUAAAAAAAAUUUAG
CUP1
psf Mutant
psf
UACUAACAUCGUUUGUUGUUUCGAUUAGCUUCAUUCUUUUUGUUGCUAUAAAAAAAAUUUAG
CUP1
Out of Frame
In Frame

Figure 2. Primer Extension of Splice Site Competition RNAs in Wild Type and *psf1-1* Strains.

(A) Primer Extension +T PyDOWN RNAs from a wild type strain, *PSF1*, or mutant strain, *psf1-1*. Primer extension products corresponding to use of the branchsite proximal (MP) or branchsite distal (MD) 3' splice site, or U1 snRNA internal control (U1) are indicated.

(B-D) Primer extension of +A WT, +TA and +AT RNAs from a wild type and *psf1-1* strain. Primer extension products are labeled as in 2A.

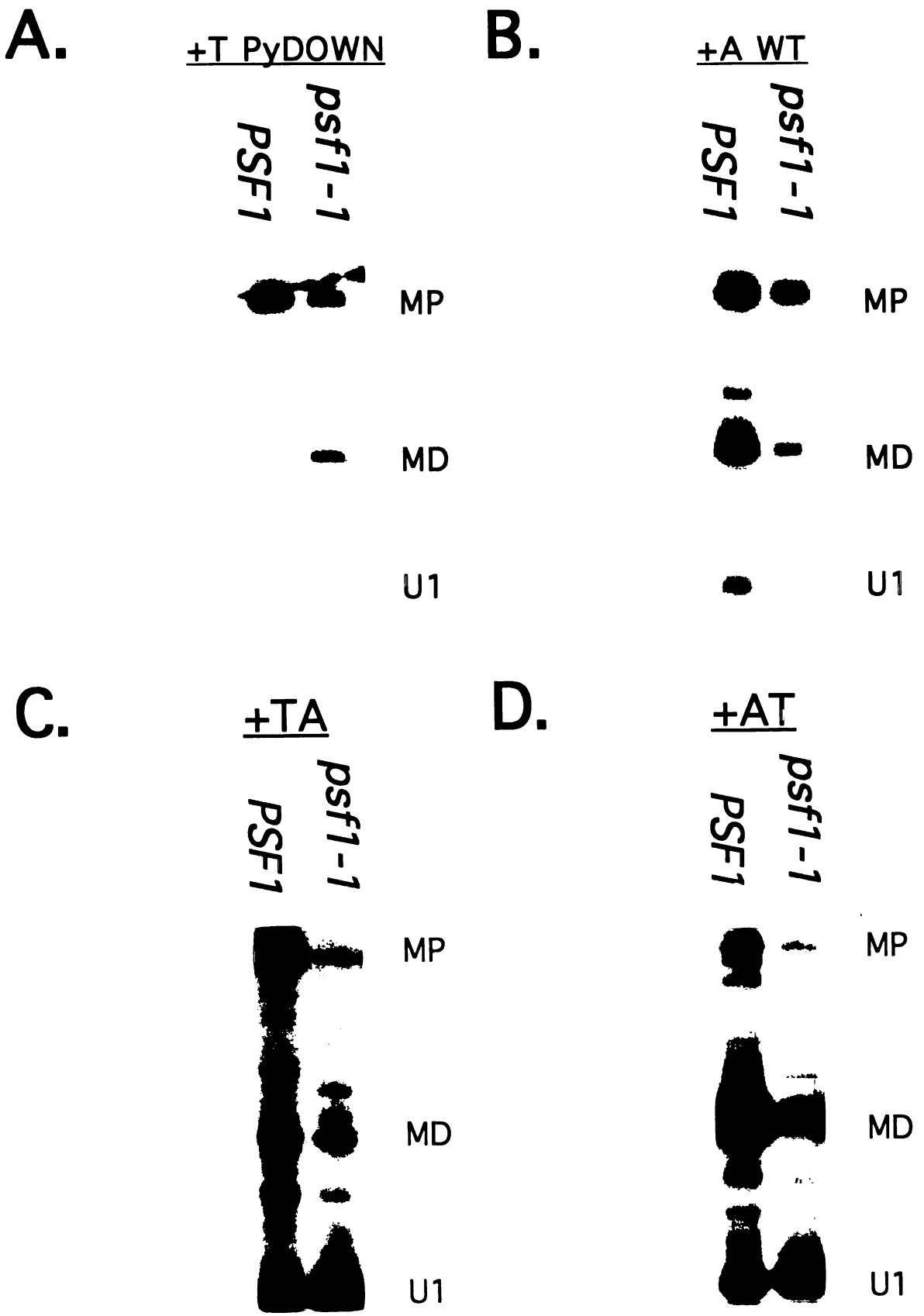
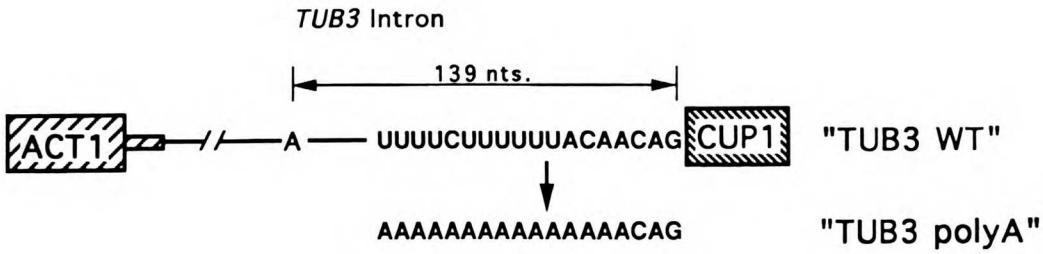

A.
+T PyDOWN
PSF1
psf1-1
MP
MD
U1
B.
+A WT
PSF1
psf1-1
MP
MD
U1
C.
+TA
PSF1
psf1-1
MP
MD
U1
D.
+AT
PSF1
psf1-1
MP
MD
U1

Figure 3. Splicing of the *TUB3* Intron in wild type and *psf1-1* Strains.

(A) Structure of the *ACT1-TUB3-CUP1* (TUB3 WT) fusion construct and TUB3 polyA fusion construct. Actin sequences are indicated by the left-hatched region, *TUB3* sequences by the lines and nucleotides and *CUP1* sequences by the right-hatched region. The sequence of the *TUB3* 3' splice site is shown.

(B) Primer extension of the *ACT1-CUP1*, TUB3 WT and TUB3 polyA RNAs in wild type, *PSF1*, and mutant, *psf1-1* , strains. The left side markers are for products of *ACT1-CUP1* splicing (lanes 1 and 2) and indicate mature message (M) and lariat intermediate (LI). A strong primer extension stop (U1*) from the *ACT1-CUP1* RNA obscures the U1 internal control in lanes 1 and 2. The right side markers are for products of TUB3 WT and TUB3 polyA splicing (lanes 3-6) and indicate mature message (M), lariat intermediate (LI) and U1 snRNA internal control (U1). Lane marker m indicates size markers which are underexposed. The *ACT1-CUP1* lariat intermediate is underexposed and not visible here. The size difference between the two mature messages is due to a ~15 nucleotide shortening of the *ACT1-CUP1* 5' untranslated region during cloning of *ACT1-TUB3-CUP1* and other constructs used in this work.

A.


TUB3 Intron
139 nts.
ACT1
A
UUUUCUUUUUUACAACAG
CUP1
"TUB3 WT"
AAAAAAAAAAAAAAACAG
"TUB3 polyA"

B.

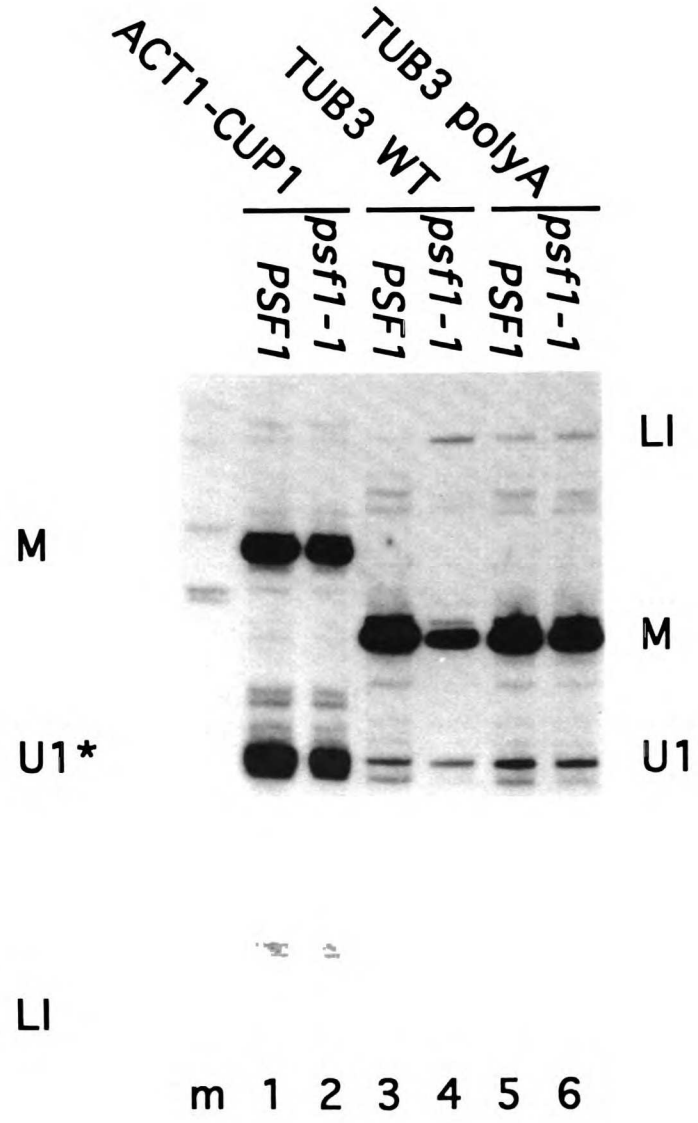

ACT1-CUP1
TUB3 WT
TUB3 polyA
PSF1
psf1-1
PSF1
psf1-1
PSF1
psf1-1
LI
M
M
U1*
U1
LI
m 1 2 3 4 5 6

Figure 4. Effect of the *psf1-1* Mutation on Splicing of Actin 3' Splice Site Derivatives.

Primer extension of wild type actin and actin 3' splice site mutants in wild type, *PSF1*, or mutant, *psf1-1*, strains. The 3' splice site mutations all reside in the final three nucleotides of the actin intron in the *ACT1-CUP1* construct and are indicated above each lane. Primer extension products are marked as in Figure 3 except L indicates lariat intermediate.

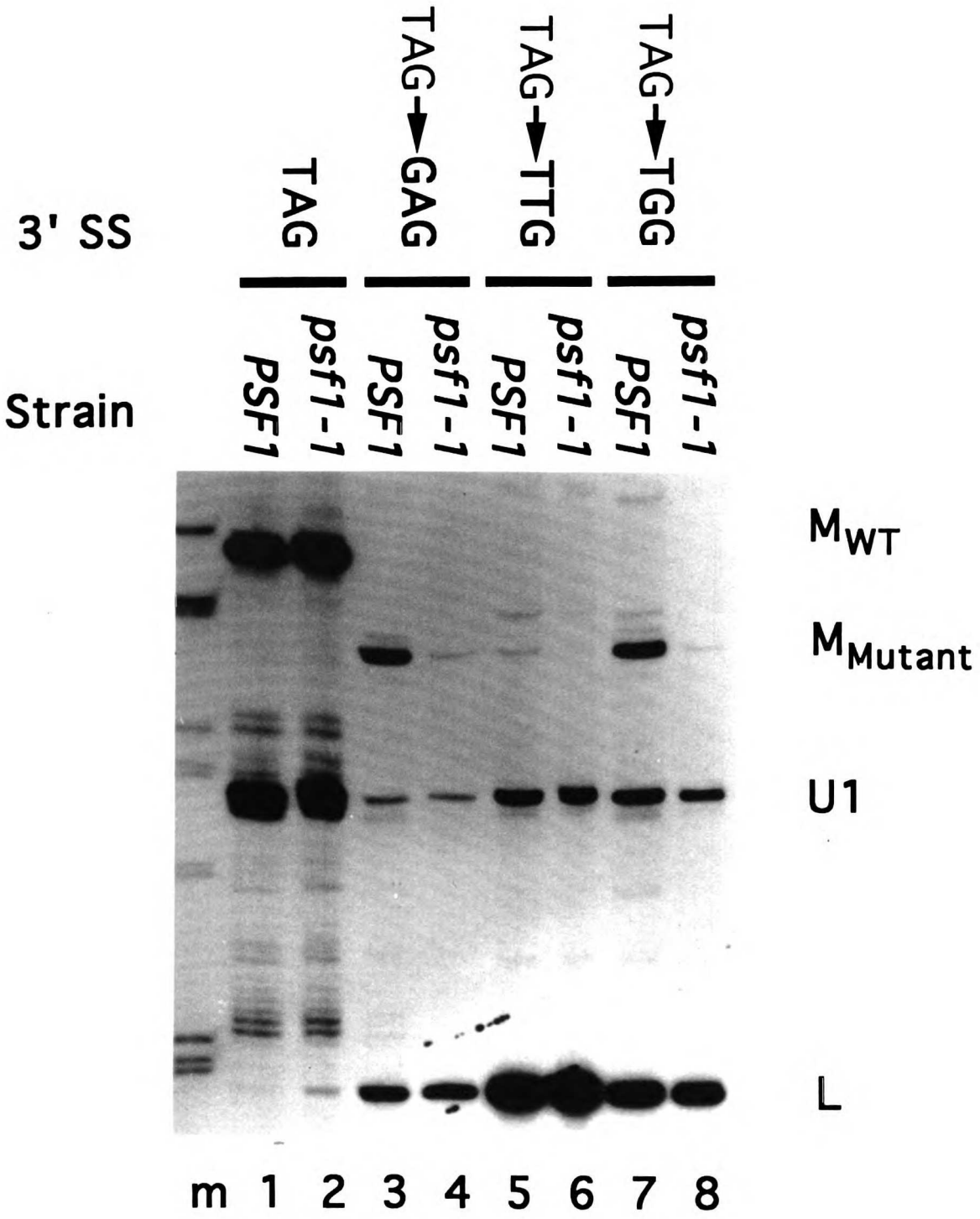

3' SS
TAG
TAG→GAG
TAG→TTG
TAG→TGG
Strain
PSF1
psf1-1
PSF1
psf1-1
PSF1
psf1-1
PSF1
psf1-1
M_WT
M_Mutant
U1
L
m 1 2 3 4 5 6 7 8

Figure 5. Allele Specificity of the *psf1-1* Phenotype.

Primer extensions of +T PyDOWN RNAs from various *prp8* strains grown at the permissive temperature of 25°. Products of primer extension are marked as in figure 2A. The *prp8* strains are described in Methods.

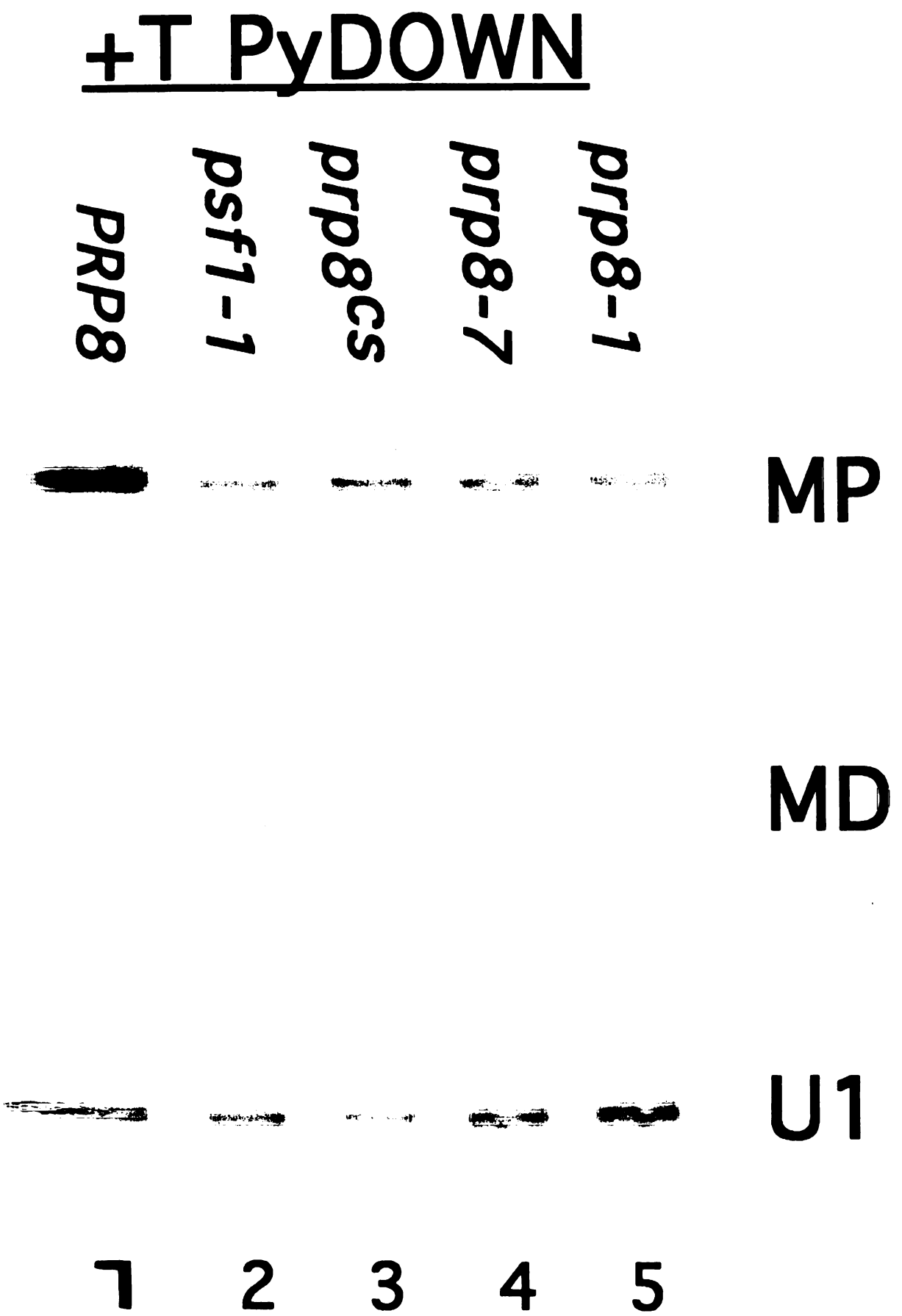

+T PyDOWN
PRP8
psf1-1
prp8cs
prp8-7
prp8-1
MP
MD
U1
1
2
3
4
5

Figure 6. UV Crosslinking of Prp8 to the 3' Splice Site

(A) Diagram of splicing substrate XL7. The 3' splice site region is shown in detail with the A to G point mutation at the 3' splice site indicated by an arrow and the single radiolabeled phosphate by a * symbol. The 15 nucleotide product of RNase T1 digestion is underlined.

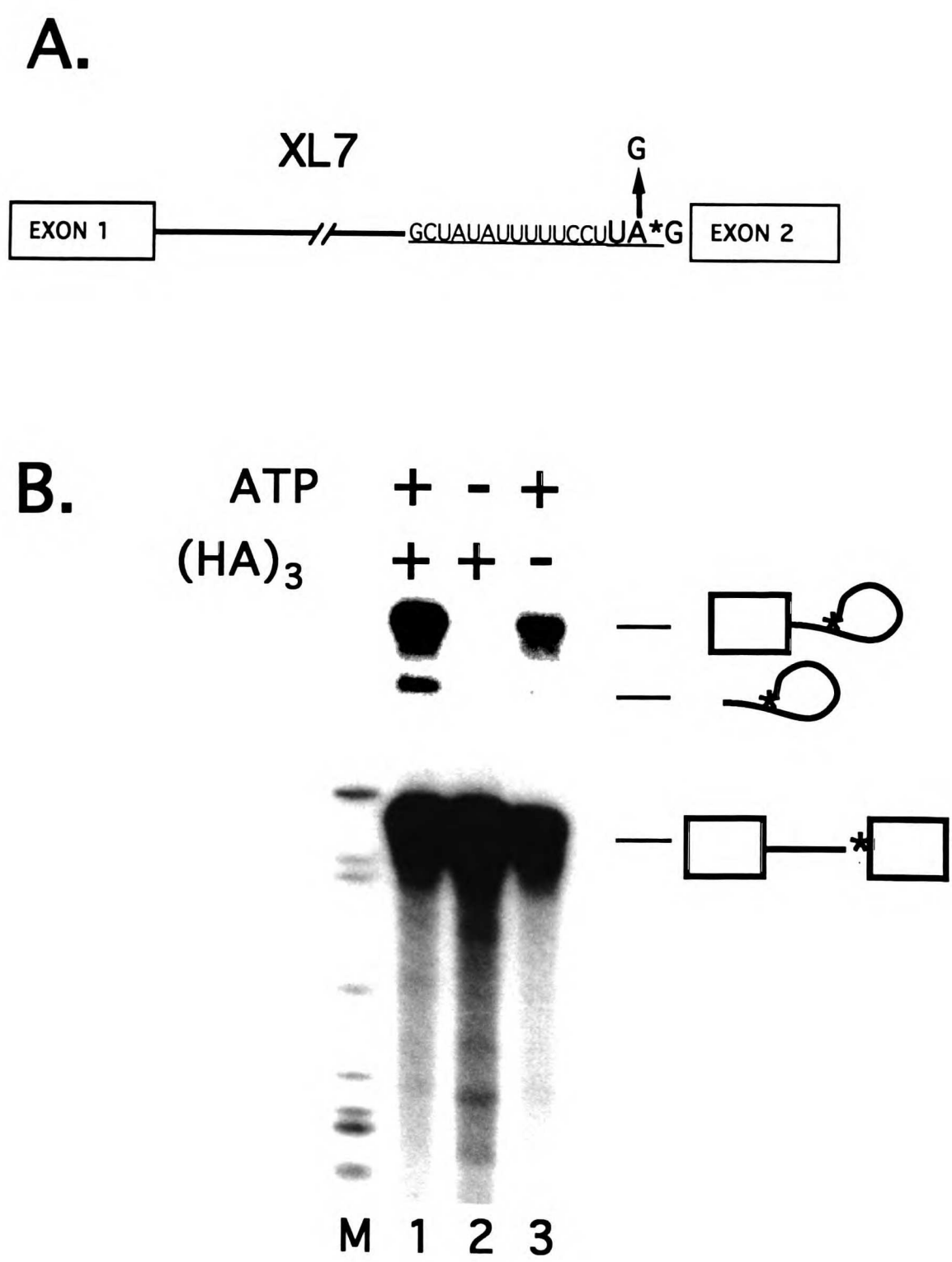
(B) Splicing of XL7 in extracts derived from PRP8 epitope tagged and untagged extracts. Presence or absence of triple hemagglutinin epitope on Prp8 is indicated by + or - symbol next to $(HA)_3$ above each lane. Presence or absence of ATP in the splicing reaction is similarly indicated. Bands corresponding to unspliced precursor, lariat intermediate and excised lariat intron are indicated by cartooned symbols next to gel with the radiolabeled phosphate molecule represented by a * symbol.

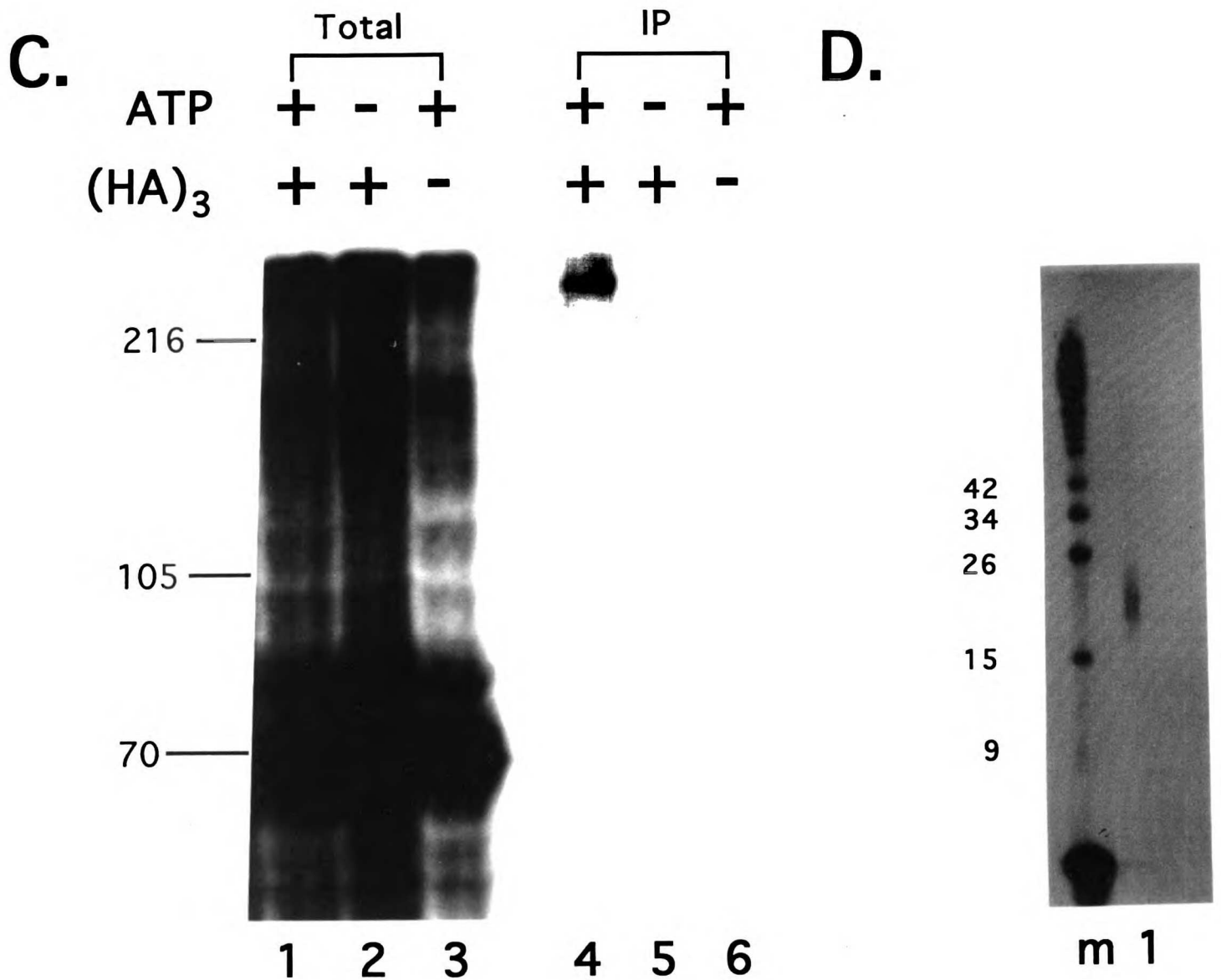
(C) UV crosslinking of samples from 6B. Lanes 1-3 represent total crosslinked proteins from 6B and lanes 4-6 represent crosslinked proteins that were immunoprecipitated with an anti-hemagglutinin epitope antibody. Markers above each lane are as in 6B.

(D) Analysis of crosslinked and immunoprecipitated RNase T1 fragment. Proteinase treated and extracted immunoprecipitate fractionated on a 20% denaturing polyacrylamide gel is in lane 1. Sizes of 5' end labeled DNA markers (HpaII digested PBR325) are indicated in lane m.

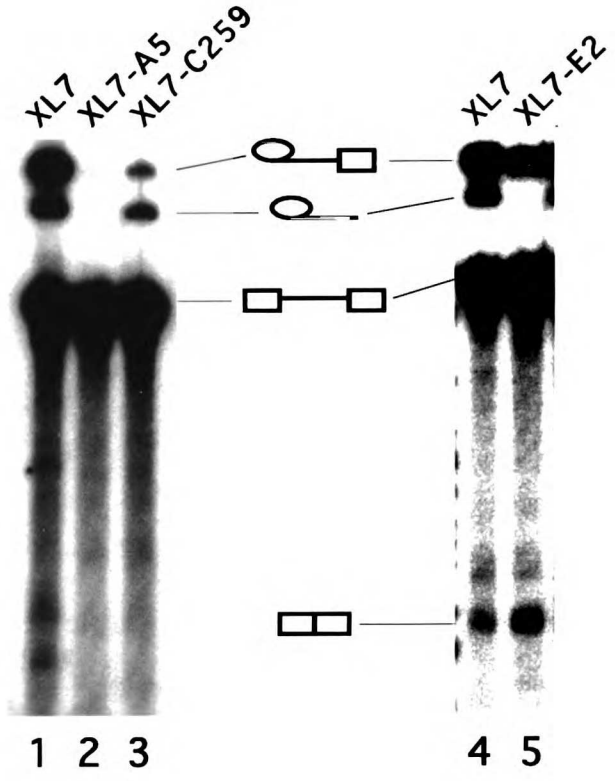
(E) Splicing of XL7 and its derivatives. Products of splicing are cartooned next to each band. Derivatives are described in the text.

(F) Crosslinking of Prp8 to XL7 and its derivatives. SDS-PAGE analysis of samples from (E) that were crosslinked and immunoprecipitated as in (C). Only the region of the gel where Prp8 migrates is shown.


A.
XL7
G
EXON 1
GCUAUAUUUUUCCUUA*G
EXON 2
B.
ATP + - +
(HA)3 + + -
M 1 2 3


C.
Total
IP
ATP + - + + - +
(HA)3 + + - + + -
216
105
70
1 2 3 4 5 6
D.
42
34
26
15
9
m 1

E.


XL7
XL7-A5
XL7-C259
XL7
XL7-E2
1 2 3
4 5

F.

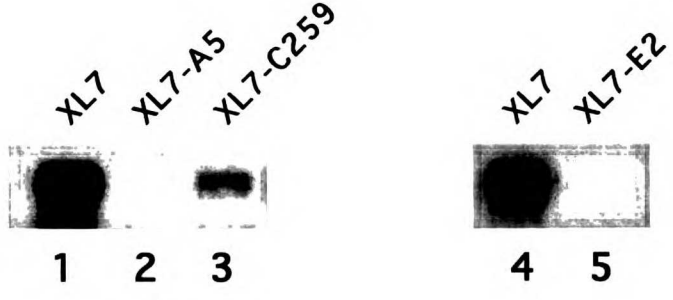

XL7
XL7-A5
XL7-C259
XL7
XL7-E2
1 2 3
4 5

Figure 7. Model for the Role of Prp8 and U5 snRNA in 3' Splice Site Selection and the Second Step of Splicing.

A splicing substrate which has formed lariat intermediate and is about to undergo the second step of splicing is cartooned. Dark lines, nucleotide sequences and darkly shaded rectangles represent the lariat structure, 3' splice site sequence and exons 1 and 2. Interactions which might serve to position exon one and the 3' splice site together are depicted. Prp8 protein is shown interacting with the uridine tract at the 3' splice site (this work) and with exon one (Wyatt et al., 1992). Prp8 is also a constituent of U5 snRNP (U5) (Lossky et al., 1987) which is depicted only in RNA form (thick dark lines) with its conserved loop nucleotides interacting with exons one and two (Newman & Norman, 1991; Newman & Norman, 1992; Wassarman & Steitz, 1992; Wyatt et al., 1992; Cortes et al., 1993; Sontheimer & Steitz, 1993).

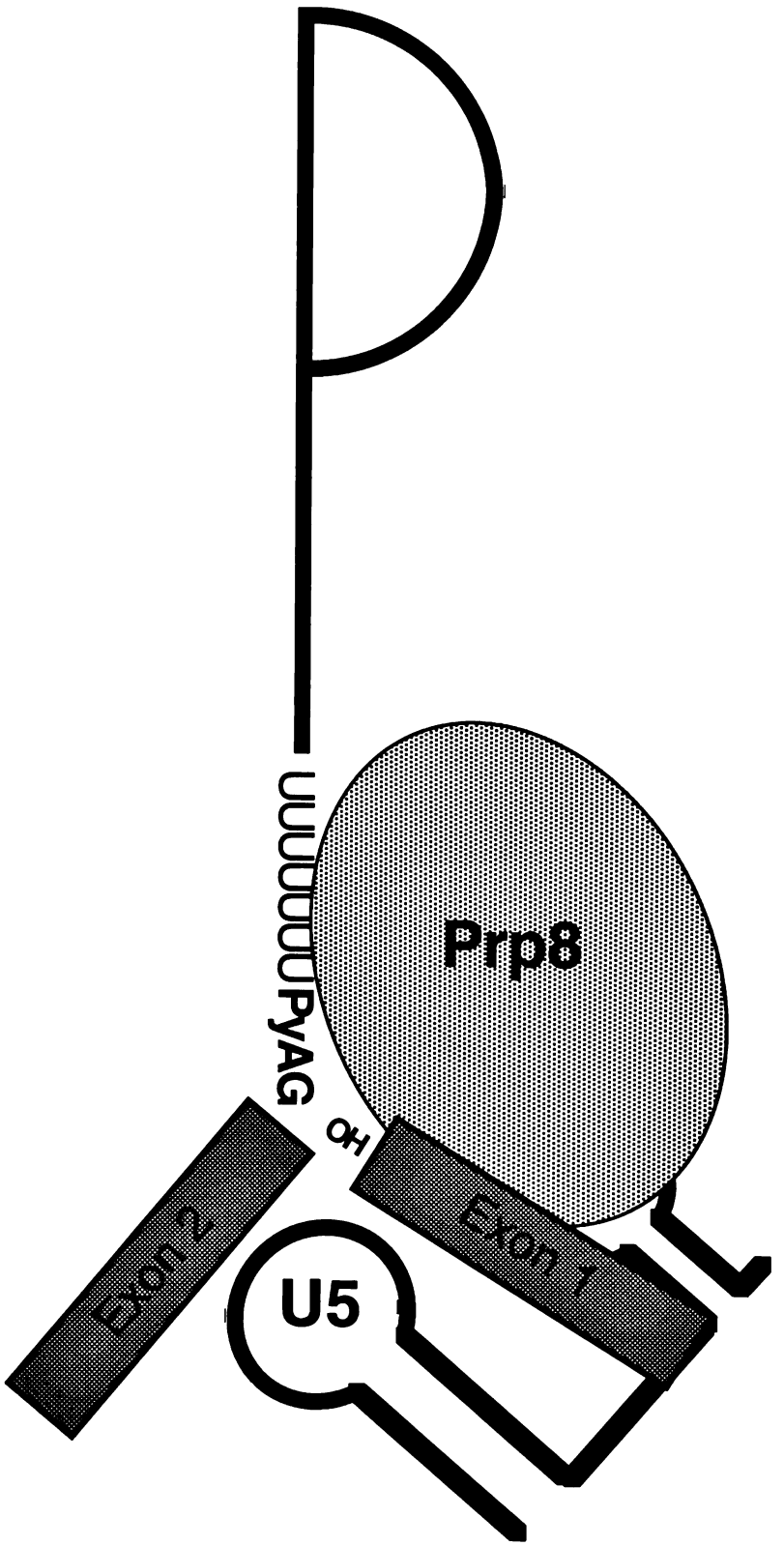

UUUUUUPyAG
Prp8
Exon 2
Exon 1
U5

Table 1.

[a] Ratio of branchpoint proximal (MP) to branchpoint distal (MD) mature message quantitated by a Molecular Dynamics PhosphorImager. This ratio reflects the relative rates of splicing to each competing splice site (see Frank and Guthrie, 1992). Data are expressed as the mean of at least three independent measurements ± standard deviation.

[b] Relative shift in the MP/MD ratio away from use of the uridine-rich 3' splice site in the mutant strain compared to wild type: $(MP/MD)_{WT}/(MP/MD)_{psf1-1}$ for +T PyDOWN and +TA; $(MD/MP)_{WT}/(MD/MP)_{psf1-1}$ for +A WT and +AT.

[c] Maximum copper resistance of strain measured with distal splice site (MD) in frame and proximal splice site (MP) out of frame. +/- indicates marginal growth at that concentration of copper. Note that copper resistance reflects only the absolute amount of (MD) mature message and not the relative ratios of splice site usage.

Table 1	Splice Site Competitions in *psf1-1* Strains: Quantitation of Splice Site Usage by Primer Extension and Copper Growth		
Strain	RNA	MP/MD[a]	Copper Resistance[b]
WT	+T PyDOWN	22 ± 0.3	.05 mM +/-
psf1-1	+T PyDOWN	1.6 ± .10	.18 mM
WT	+A WT	.86 ± .04	1 mM
psf1-1	+A WT	3.1 ± .03	.1 mM
psf1-1/WT[c]	+T PyDOWN	2.9 ± .3	.1 mM

Table 2.

[a] Ratio of mature message (M) to lariat intermediate (L) quantitated by a Molecular Dynamics Phosphorimager. M/L provides a measure of the overall efficiency of the second catalytic step of splicing (see Frank and Guthrie, 1992). Data are expressed as the mean of at least three independent measurements ± standard deviation. n.a. indicates not applicable.

[b] Maximum copper resistance of a strain measures absolute amounts of mature message produced. Note that copper growth values (M only) will not always correspond to those for M/L. For discussion of this, see (Pikielny & Rosbash, 1985). Also, copper resistance is not entirely linear with respect to the amount of message (Lesser & Guthrie, 1993).

[c] Copper resistance values below .025 mM are indistinguishable from a *CUP1* null mutant which grows at .013 mM copper.

Table 2 Quantitation of Splicing of Wild Type *ACT1-CUP1* and *ACT1-TUB3-CUP1* Derivatives by Primer Extension and Copper Growth.

Strain	RNA	M/L[a]	Copper Resistance[b]
WT	ACT1-CUP1	338 ± 31	2 mM
psf1-1	ACT1-CUP1	110 ± 3	1.5 mM
WT	TUB3 WT	203 ± 2.0	2 mM
psf1-1	TUB3 WT	10.2 ±. 30	.25 mM
WT	TUB3 polyA	62 ± 5.8	1 mM
psf1-1	TUB3 polyA	41 ± .75	1 mM
WT	none[c]	na	.013 mM
psf1-1	none	na	.013 mM

Table 3 Quantitation of Splicing to Wild Type *ACT1-CUP1* and *ACT1-CUP1* 3' Splice Site Derivatives by Primer Extension and Copper Growth.

Strain	RNA	M/L[a]	Copper Resistance[b]
WT	ACT1-CUP1 WT	338±31	2 mM
psf1-1	ACT1-CUP1 WT	110±13	1.5 mM
WT	ACT1-CUP1 3' GAG	.65±.04	.1 mM
psf1-1	ACT1-CUP1 3' GAG	.056±.002	.013 mM
WT	ACT1-CUP1 3' TTG	.0054±.0019	.013 mM
psf1-1	ACT1-CUP1 3' TTG	.00085±.0005	.013 mM
WT	ACT1-CUP1 3' TGG	.18±0.00	.025 mM
psf1-1	ACT1-CUP1 3' TGG	.014±.002	.013mM
WT	none[c]	na	.013mM
psf1-1	none	na	.013mM

CHAPTER 2

Prp16p, Slu7p and Prp8p Interact With the 3' Splice Site in Two Distinct Stages During the Second Catalytic Step of pre-mRNA Splicing

ABSTRACT

For the second catalytic step of pre-mRNA splicing to occur, a 3' splice site must be selected and juxtaposed with the 5' exon. Four proteins, Prp16p, Slu7p, Prp17p, Prp18p, and an integral spliceosomal protein, Prp8p, are known to be required for the second catalytic step. *prp8-101*, an allele of *PRP8* defective in 3' splice site recognition, exhibits specific genetic interactions with mutant alleles of the other second step splicing factors. The *prp8-101* mutation also results in decreased crosslinking of Prp8p to the 3' splice site. To determine the role of the step-two specific proteins in 3' splice site recognition and in binding of Prp8p to the 3' splice site, we performed crosslinking studies in mutant and immunodepleted extracts. Our results suggest an ordered pathway in which, after the first catalytic step, Prp16p crosslinks strongly to the 3' splice site and Prp8p and Slu7p crosslink weakly. ATP hydrolysis by Prp16p affects a conformational change that reduces the crosslinking of Prp16p with the 3' splice site and allows stronger crosslinking of Prp8p and Slu7p. Thus, the 3' splice site appears to be recognized in two stages during the second step of splicing. Strong 3' splice site crosslinking of Prp8p and Slu7p also requires the functions of Prp17p and Prp18p. Therefore, Prp8p and Slu7p interact with the 3' splice site at the latest stage of splicing prior to the second catalytic step that can currently be defined and may be at the active site.

INTRODUCTION

Nuclear pre-mRNA splicing involves the recognition and removal of introns from messenger RNA precursors. Five small ribonucleoproteins (U1, U2, U4, U5 and U6 snRNPs), together with multiple accessory proteins, recognize and assemble onto intron-containing RNAs in an ordered fashion to form the spliceosome and carry out the two transesterification reactions that characterize pre-mRNA splicing (see below for reviews). In the first chemical step, the 2' hydroxyl of an internal (branchsite) adenosine attacks the 5' splice site phosphodiester bond to generate the lariat intermediate and free 5' exon; in the second chemical step, the 3' hydroxyl of the free 5' exon attacks the 3' splice site phosphodiester bond forming ligated exons and an excised lariat intron.

These reactions require accurate identification and juxtaposition of splice sites. While a great deal has been learned about this process for the first catalytic step of splicing, much less is known about 3' splice site selection and the second catalytic step (Green, 1991; Guthrie, 1991; Rymond and Rosbash, 1992; Moore et al., 1993; Madhani and Guthrie, 1994a). 3' splice sites in most organisms contain a pyrimidine-rich tract upstream of an invariant AG dinucleotide at the 3' splice junction. Moreover, a specific interaction between the first and last guanosine residues in introns plays an important functional role in 3' splice site utilization (Parker and Siliciano, 1993). In mammals, the pyrimidine tract is recognized by a factor, U2AF, that is required at an early stage in spliceosome assembly (Zamore and Green, 1991). Later, the pyrimidine tract is bound by PSF(PTB associated splicing factor) which is required for the second catalytic step (Gozani et al., 1994). In both *Saccharomyces cerevisiae* and mammals, U5 snRNA has been shown to base pair with exon sequences adjacent to the 3' splice site; however, the lack of sequence conservation in these sequences suggests that this interaction does not normally play a major determinative role in 3' splice site selection (Newman and Norman, 1991; Newman and Norman, 1992; Wyatt et al., 1992; Cortes et al., 1993; Sontheimer and Steitz, 1993; Madhani and Guthrie, 1994a).

The U5 snRNP-associated protein Prp8p was originally identified from a temperature sensitive mutant allele, *prp8-1*, that blocks splicing prior to the first catalytic step *in vitro* and *in vivo* (Lossky et al., 1987; Jackson et al., 1988; Brown and Beggs, 1992). Recently, however, we identified a novel allele, *prp8-101*, that is specifically defective for 3' splice site selection. This allele impairs recognition of the uridine tract preceding the 3' splice site and defines a function for Prp8p at the second catalytic step of splicing *in vivo* (Umen and Guthrie, 1995). Site-specific UV crosslinking suggests that Prp8p mediates 3' splice site selection through direct binding to the 3' splice site (Teigelkamp et al., 1995a; Umen and Guthrie, 1995).

In addition to the factors mentioned above, genetic screens in *Saccharomyces cerevisiae* have identified four proteins that are required specifically for the second catalytic step: Prp16p, Slu7p, Prp17p and Prp18p. These proteins are "exchangeable" since they can be removed from *in vitro* splicing reactions and then added back to complement pre-assembled spliceosomes (Vijayraghavan and Abelson, 1990; Schwer and Guthrie, 1991; Horowitz and Abelson, 1993a; Athar and Schwer, 1995; Jones et al., 1995). Additionally, *PRP16*, *PRP17*, *PRP18*, and *SLU7* share a unique set of genetic interactions with each other but not with other splicing factors, suggesting a physical or functional association of the encoded proteins (Frank et al., 1992; Jones et al., 1995).

Originally identified as a suppressor of a branchsite mutation, *PRP16* belongs to a family of RNA-dependent ATPases, the so called DExH box proteins (Couto et al., 1987; Burgess et al., 1990; Schmid and Linder, 1992). Prp16p binds to the spliceosome after the first catalytic step of splicing; it then promotes an ATP-dependent conformational change that leads to protection of the 3' splice site from oligonucleotide-directed RNase H cleavage (Schwer and Guthrie, 1991; Schwer and Guthrie, 1992a). Whether this protection is due to Prp16p or other splicing factors is not known.

SLU7 was identified in a screen for mutants that are synthetically lethal with U5 snRNA (Frank et al., 1992). Although Slu7p does not appear to be stably associated with

U5 or other snRNPs (D. N. Frank, unpublished observations), the protein contains a sequence motif ($CX_2CX_4HX_4C$), termed a "zinc knuckle", that is implicated in retroviral RNA binding (Frank and Guthrie, 1992 and references therein). Interestingly, an allele of *SLU7*, *slu7-1*, affects 3' splice site selection, suggesting a possible RNA binding site for the protein (Frank and Guthrie, 1992).

Prp17p and Prp18p are encoded by non-essential genes. Absence of either causes a partial block to the second step of splicing *in vitro* (Vijayraghavan and Abelson, 1990; Frank et al., 1992; Horowitz and Abelson, 1993b; Jones et al., 1995). Whereas Prp16p and Prp17p act at or prior to an ATP-requiring stage of splicing (Schwer and Guthrie, 1991; Jones et al., 1995), Prp18p and Slu7p do not require ATP to promote the second catalytic step (Horowitz and Abelson, 1993a; Athar and Schwer, 1995; Jones et al., 1995). Thus Prp16p/Prp17p and Prp18p/Slu7p define at least two stages in the second step of splicing, one ATP-dependent and one ATP-independent.

Here we investigate the *in vitro* phenotype of *prp8-101*. We find that the mutant protein causes a block to the second catalytic step of splicing and is impaired in 3' splice site crosslinking. The *prp8-101* mutant also displays specific genetic interactions with alleles of *PRP16*, *SLU7*, *PRP17* and *PRP18*, establishing a functional relationship between Prp8p and these second step splicing factors. We utilize site-specific UV crosslinking to test whether the exchangeable second step splicing factors interact with the 3' splice site, and to determine the timing of these interactions relative to that of Prp8p. We find that the 3' splice site is recognized in at least two distinct stages. The first is characterized by strong 3' splice site crosslinking of Prp16p and weaker crosslinking of Slu7p and Prp8p. After hydrolysis of ATP by Prp16p, Prp8p and Slu7p crosslinking to the 3' splice site increases while Prp16p crosslinking is diminished. Strong 3' splice site crosslinking of Prp8p and Slu7p also requires the functions of Prp17p and Prp18p. Thus, Prp8p and Slu7p interact with the 3' splice site at the closest stage of splicing prior to catalysis that can currently be defined.

RESULTS

In vitro analysis of the *prp8-101* mutant

Previously we demonstrated that the *prp8-101* allele of *PRP8* causes a specific defect in 3' splice site uridine tract recognition (Umen and Guthrie, 1995). Furthermore, we utilized a site-specifically labeled *in vitro* splicing substrate (XL7) and UV crosslinking to show that wild type Prp8p crosslinks to the 3' splice site in active spliceosomes (Figure 1) (Umen and Guthrie, 1995). XL7 contains a point mutation (AG to GG) at the 3' splice site that slows the kinetics of the second step and enhances crosslinking of Prp8p. We wished to determine whether the *prp8-101* allele causes a block to the second step of splicing *in vitro* and how the Prp8p-3' splice site interaction is affected in the mutant strain. To test this, we prepared splicing extract from an epitope-tagged *prp8-101* strain (see MATERIALS AND METHODS) and examined splicing and 3' splice site crosslinking. In the *prp8-101* mutant extract, XL7 undergoes the first step of splicing at a similar efficiency compared with the wild type extract. However, there is a block to the second catalytic step (Figure 2A). This block is partial when the 3' splice site dinucleotide is wild type (AG) (data not shown). We consistently observed a 2- to 4-fold reduction in Prp8p-3' splice site crosslinking with the mutant *prp8-101* extract (Figure 2B). Immunoblot analysis of the crosslinked samples indicated that similar quantities of protein were immunoprecipitated in each reaction (Figure 2C). Furthermore, native immunoprecipitation of spliceosomes from wild type and mutant extracts revealed that similar amounts of Prp8p are associated with precursor and lariat intermediate (Figure 2D, lanes 2 and 3).

Interestingly, in contrast to previously reported results (Whittaker et al., 1990; Teigelkamp et al., 1995b), we detect very little association of Prp8p with excised lariat in this experiment. Even upon long exposure of the autoradiograph in Figure 2D, the proportion of excised lariat to lariat intermediate is greatly reduced in the immunoprecipitate versus the total splicing reaction (Figure 2A, lane 1 versus Figure 2D, lane 2). This may

be due to differences between the two splicing substrates utilized, a difference in epitope accessibility in complexes containing excised lariat versus lariat intermediate and precursor, or differences in spliceosome disassembly rates in the respective splicing extracts. To test the second possibility, we crosslinked Prp8p to XL7 in a splicing reaction and analyzed Prp8p-XL7 crosslinking by denaturing immunoprecipitation (see MATERIALS AND METHODS) (Figure 2E). This experiment yields a similar profile of precipitated RNA species as the native immunoprecipitation. We conclude that in our extracts, Prp8p is primarily associated with precursor and lariat intermediate and only a small fraction could be associated with excised lariat. The discrepancy with previous results is either due to differences in the two substrates that were utilized or to differences in spliceosome disassembly rates. Our result further suggests that the Prp8p-3' splice site interaction is altered or destabilized after the second catalytic step.

Genetic interactions between *PRP8* and other second step splicing mutants.

Synthetic lethal analysis has proven useful in *de novo* identification of interacting splicing factors and in functionally grouping known splicing factors (Frank et al., 1992; Liao et al., 1993; Ruby et al., 1993; Wells and Ares, 1994). Mutant alleles of the second step splicing proteins encoded by *PRP16*, *SLU7*, *PRP17* and *PRP18* exhibit synthetic lethality with each other but not with mutant alleles of proteins required for the first step of splicing (Frank et al., 1992). Since the *prp8-101* allele behaves specifically as a second step mutant both *in vivo* and *in vitro*, we tested whether it interacts genetically with mutant alleles of second step splicing genes. In principle, just as first step splicing factors fall into different genetic sub-groups based on common biochemical functions, the same could be true of second step splicing factors; that is, *prp8-101* might define its own sub-group, or it might be related in function to the other second step factors.

We employed a novel method for constructing and analyzing double mutants of *prp8-101* or *prp8-1* and alleles of both first and second step splicing mutants (Figures 3A

and 3B). This method involves the direct construction and analysis of double mutant strains via integrative transformation rather than the more laborious indirect analysis of meiotic progeny. It is particularly useful for rapidly screening one or two mutants against a large set of test strains. Either *prp8-101* or *prp8-1* were introduced into test strains carrying a chromosomal or plasmid-borne *prpx* mutation and a *URA3* marked plasmid bearing a wild type copy of the same *PRPX* gene. Selection against the *URA3-PRPX* plasmid with the drug 5-fluoro-orotic acid (5FOA) yields a double mutant *prp8 prpx* strain (Figure 3B). Failure to grow on 5FOA indicates that the double mutant strain is inviable and that the mutations are synthetically lethal. Compared with *prp8-101* strains, which grow at near wild type rates and are mildly temperature sensitive, *prp8-1* strains grow slowly and are strongly temperature sensitive. Therefore, *prp8-1* served as a good control for specificity in these experiments.

We analyzed three first step splicing mutants (*prp2-1*, *prp3-1* and *prp24-6*) and at least one allele of all the second step splicing factors (*prp16-101*, *prp16-2*, *prp16-301*, *prp17-1*, *prp17-2*, *prp18-1*, *slu7-1* and *slu7-ccss* [double mutation in the "zinc knuckle"(Frank and Guthrie, 1992)]). Strikingly, we found that *prp8-101* is synthetically lethal with at least one allele of all the second step splicing factors but displays no genetic interactions with the first step mutants (Figure 3C). For *PRP16* we only observed synthetic lethality with a cold sensitive allele (*prp16-301*) but not with a temperature sensitive allele (*prp16-2*) or a mutant branchsite suppressor (*prp16-101*). In contrast, *prp8-1* showed no genetic interactions with the *prp* mutants we tested. Prp8p, therefore, appears to function coordinately with the other second step splicing proteins.

Crosslinking of Prp16p and Slu7p to the 3' splice site

To test whether second step proteins besides Prp8p might function in 3' splice site recognition, we determined whether any can be crosslinked to the 3' splice site. Prp16p and Slu7p are particularly attractive candidates since their sequences and/or biochemical

properties predict that they interact with spliceosomal RNAs (Schwer and Guthrie, 1991; Frank and Guthrie, 1992). Moreover, Slu7p is known to participate in 3' splice site selection (Frank and Guthrie, 1992). We utilized our previously characterized XL7 substrate and crosslinking assay (Figure 1B) to determine whether Prp16p, Slu7p, Prp17p or Prp18p crosslink to the 3' splice site. We also tested the U1 70K protein (Snp1p) as a control for a splicing factor most likely involved in the first step of splicing. After splicing and crosslinking in a wild type extract, antiserum raised against each of these proteins was used for immunoprecipitation followed by denaturing PAGE and autoradiography. Prp16p and Slu7p crosslinked to the 3' splice site in this assay whereas Prp17p, Prp18p and Snp1p did not (Figure 4B).

As controls for the immunoprecipitations, we looked at immunoprecipitation of U1 snRNA by αSnp1p serum or U4, U5 and U6 snRNAs by αPrp18 serum under conditions similar to those used in the crosslinking experiments (Figure 4C, MATERIALS AND METHODS). Under our experimental conditions, we were able to immunopreciptate U1 snRNA with αSnp1p serum but could not precipitate U4, U5 and U6 with αPrp18p serum (Figure 4C, lanes 2 and 3). However, we could immunoprecipitate U4, U5 and U6 with antibodies specific for the epitope-tagged Prp8p (Figure 4C, lane 1). Note that some U6 snRNA also precipitated non-specifically in this experiment. Since Prp17p is often obscured on immunoblots by IgG after immunoprecipitation, we compared the supernatant from the αPrp17p immunoprecipitation experiment to a similar amount of undepleted splicing extract. Comparison of lanes 1 and 2 in Figure 4D shows that a substantial fraction of Prp17p was depleted in the immunoprecipitation after crosslinking. In summary, these controls indicate that only a specific subset of splicing factors crosslink to the 3' splice site.

To determine whether Prp16p and Slu7p crosslink to the 3' splice site in active spliceosomes, we employed two derivatives of XL7, XL7-A5 and XL7-C259 (Umen and Guthrie, 1995). XL7-A5 contains a G to A mutation at the fifth position of the intron and

undergoes no splicing. We see no 3' splice site crosslinking to either Prp16p or Slu7p with this substrate (Figures 4A and 4E, lane 1 versus 2). XL7-C259 contains an A to C mutation at the branch residue and splices 5- to 10- fold less efficiently than XL7 (Figure 4A lane 1 versus 3). Similarly, Prp16p and Slu7p show a 5- to 10- fold decrease in 3' splice site crosslinking with XL7-C259 (Figure 4E, lane 1 versus 3). These results indicate that Prp16p and Slu7p crosslink to the 3' splice site in active spliceosomes.

To ascertain whether the Prp16p and Slu7p crosslinks were specific to the 3' splice site, we used a modified crosslinking substrate (XL7-E2). In XL7-E2, the second exon is uniformly labeled beginning nine nucleotides downstream of the 3' splice junction and is substituted with 5-bromo-uridine (Umen and Guthrie, 1995). Both of these proteins failed to crosslink to XL7-E2 (Figure 4A, lane 5; Figure 4E, lanes 4 and 5). Immunoblotting of the samples indicated that the crosslinked protein in each immunoprecipitate comigrates with the signal from the immunoblot (data not shown), and that similar amounts of protein were immunoprecipitated in each sample (Figure 4F, lane 1 versus 2 and lane 3 versus 4). Therefore, Prp16p and Slu7p interact specifically with the 3' splice site.

Timing and Requirements for 3' Splice Site Crosslinking by Prp8p, Prp16p and Slu7p

Establishing the relative timing of 3' splice site interaction for Prp8p, Prp16p and Slu7p is critical for understanding their roles in the second step of splicing. In particular, it is important to determine the time of 3' splice site binding relative to the catalytic steps of splicing. Furthermore, the relative order of 3' splice site interaction for these proteins with respect to each other may be informative regarding their precise biochemical functions. To address these questions, we used mutant or immunodepleted extracts that were blocked at various stages of the splicing reaction. A decrease or loss of crosslinking in the mutant extract compared to the control indicates that the activity of the missing or mutant protein is required prior to the crosslink. Conversely, no change or an increase in crosslinking in the

mutant versus the control extract indicates that the activity of the mutant or missing protein is not required for the crosslink.

We began our analysis by examining crosslinking in extracts prepared from a *prp2-1* strain. Prp2p function is required after complete spliceosome assembly and just prior to the first catalytic step (Cheng and Abelson, 1987; Kim and Lin, 1993). Since the *prp2-1* encoded protein is thermolabile *in vitro*, we compared crosslinking of Prp8p, Prp16p and Slu7p in *prp2-1* extracts with and without heat inactivation. These extracts were fully complementable with purified Prp2p (data not shown). We found that crosslinking is abolished for all three proteins when spliceosomes are blocked at the Prp2p step (Figures 5A and 5B). Thus, crosslinking appears to occur either immediately prior to or after the first catalytic step of splicing.

Next, we examined the Prp16p-dependent step of splicing. Spliceosomes were formed in extracts that had been immunodepleted for Prp16p, and then glucose and hexokinase were added to deplete ATP. We added either buffer, ATP, purified Prp16p, or both purified Prp16p and ATP to these spliceosomes and continued the incubation to allow completion of splicing. Splicing and crosslinking were then examined in each sample. As previously reported, the splicing reaction can only be complemented with the addition of both ATP and Prp16p (Figure 5C) (Schwer and Guthrie, 1991). Furthermore, Prp8p and Slu7p crosslink to the 3' splice site in the absence of Prp16p but require hydrolysis of ATP by Prp16p for maximal crosslinking (Figure 5D, lanes 1-3 versus lane 4). We consistently observe a 2- to 3-fold increase in Prp8p and Slu7p crosslinking to the 3' splice site when the reaction is complemented with Prp16p and ATP. In contrast, Prp16p crosslinks to the 3' splice site 2-to 3-fold more strongly in the absence of ATP than in its presence, consistent with its release from spliceosomes upon ATP hydrolysis (Figure 5D, lane 3 versus 4) (Schwer and Guthrie, 1991). The slowed reaction kinetics of the splicing substrate XL7 most likely account for the remaining Prp16p in spliceosomes after complementation.

We also examined crosslinking in an extract derived from a *prp16-1* strain. The protein encoded by the *prp16-1* allele can bind spliceosomes but is strongly reduced for ATP hydrolysis and does not release from spliceosomes. The mutant protein can thus act as a "dominant negative" for the second catalytic step (Figure 5E) (Schwer and Guthrie, 1992b). The ATPase defect caused by the *prp16-1* mutation might be due to poor binding of an RNA or due to another defect in the ATP hydrolysis cycle (e.g. activation of ATP hydrolysis upon RNA binding, or release after ATP hydrolysis). If *prp16-1* causes an RNA binding defect, then we might expect to see reduced Prp16p-3' splice site crosslinking in the mutant strain. In contrast, we found that the *prp16-1* mutation caused a 3- to 4-fold increase in the Prp16p-3' splice site crosslink (Figure 5F). Interestingly Prp8p- and Slu7p-3' splice site crosslinking are nearly eliminated in the *prp16-1* extract (Figure 5F). By trapping Prp16p in a state where it is bound to the 3' splice site but cannot complete the ATP hydrolysis cycle, the *prp16-1* mutant appears to prevent the binding of Prp8p and Slu7p.

To examine whether the functions of Prp17p and Prp18p are required for Prp8p, Prp16p and Slu7p to bind the 3' splice site, we utilized extracts derived from *prp17-1* and *prp18-1* mutant strains. Although these extracts cause a constitutive partial block to the second step of splicing for a wild type splicing substrate, they cause a stronger block with XL7 (Figures 5G and 5I, lane 1; data not shown). Splicing and crosslinking were examined in these extracts which were incubated with XL7 to allow spliceosome assembly and the first catalytic step to occur. At this point, excess cold competitor RNA was added to prevent further initiation of splicing, followed by a "chase" of wild type complementing extract or additional mutant extract. For both *prp17-1* and *prp18-1* extracts, we observed a modest but reproducible increase in Prp8p- and Slu7p-3' splice site crosslinking when the defect was complemented (Figures 5H and 5J, lanes 1 and 2). In contrast, the Prp16p-3' splice crosslink stays the same or is diminished slightly upon complementation. Thus, while Prp17p and Prp18p are required for Prp8p and Slu7p to crosslink maximally to the

3' splice site, their functions are not required for the Prp16p-3' splice site interaction. Moreover, Prp18p may facilitate the release of Prp16p from spliceosomes since there is usually greater Prp16p crosslinking in the mutant versus complemented reaction (Figure 5J, lanes 1 and 2).

Finally, we determined whether Prp8p or Prp16p require Slu7p to bind the 3' splice site. We followed a similar protocol as described for *prp17-1* and *prp18-1* extracts except that we utilized an extract that had been immunodepleted with αSlu7p serum (ΔSlu7). XL7 was incubated in this extract, and splicing and 3' splice site crosslinking were examined. The extent of Slu7p immunodepletion is variable and in this experiment caused only a partial second step splicing block (Figure 5K, lane 1 versus lane 3). In ΔSlu7 extracts, Prp8p-3' splice site crosslinking is consistently reduced compared to the complemented control whereas Prp16p-3' splice site crosslinking is enhanced (Figure 5L, lanes 1 and 2). Thus, similar to Prp17p and Prp18p, Slu7p is required for Prp8p to interact maximally with the 3' splice site and is also required for Prp16p to release from the 3' splice site. We have obtained similar results using *slu7-1* extracts for these experiments (data not shown). The results of these crosslinking experiments are summarized in Table 1 and Figure 6.

DISCUSSION

prp8-101 causes a 3' splice site binding defect

Using extracts derived from a *prp8-101* strain, we have demonstrated that the *prp8-101* mutation causes a second step splicing block *in vitro* and a partial reduction in Prp8p-3' splice site crosslinking. The reduction in crosslinking in the *prp8-101* extract is not a secondary result of the mutant Prp8p being destabilized from spliceosomes, as it is associated with precursor and lariat intermediate as stably as the wild type protein under our assay conditions. Since Prp8p crosslinking to the 3' splice site only occurs after the first catalytic step (Figure 5A and 5B) and since Prp8p is not associated with or crosslinked to

excised lariat in our extracts, the Prp8p-3' splice site interaction must take place in lariat intermediates prior to or during the second catalytic step. Therefore, the defect caused by this mutation is associated with a reduced interaction with the 3' splice site in spliceosomes that contain the lariat intermediate.

The partial reduction in crosslinking versus the complete block to the second catalytic step of splicing with this mutant has several possible explanations. One is that a minimal occupancy time by Prp8p at the 3' splice site is required for catalysis and the *prp8-101* mutation reduces binding below that minimum but does not eliminate binding. It is important to bear in mind that crosslinking represents an interaction between two molecules but is not a direct measure of binding. Crosslinking can overrepresent binding site occupancy of a protein since it permanently "captures" binding events in a given interval of time. Thus, the apparent modest binding defect caused by the *prp8-101* mutation may be more severe than apparent from crosslinking. This qualification also applies to the crosslinking results we obtain for Prp8p, Prp16p and Slu7p in mutant or depleted extracts (Figure 5). The modest but reproducible changes in 3' splice site crosslinking for Prp16p, Prp8p and Slu7p may represent more severe alterations in 3' splice site binding than is apparent from differences in crosslinking efficiency. A second explanation for the small reduction in Prp8p-3' splice crosslinking in *prp8-101* extracts is that reduced binding might cause only part of the splicing defect. For example, the block to splicing could involve a loss of communication between Prp8p and another splicing component that activates 3' splice site usage (see below) or could be due to binding in an altered conformation. These issues can be more readily addressed when a direct assay for 3' splice site binding is developed. We are also currently mapping the domain(s) of Prp8p that is required for its 3' splice site interaction.

prp8-101 interacts genetically with second step mutants.

By constructing double mutant strains, we have shown synthetic lethal interactions

between *prp8-101* and alleles of *PRP16*, *PRP17*, *PRP18* and *SLU7* but not with alleles of first-step splicing mutants. That a more severe allele of *PRP8*, *prp8-1* does not show genetic interactions with the second-step splicing mutants demonstrates the sensitivity and selectivity of this assay. Furthermore, among the second step splicing mutants there is some specificity; *prp8-101* is only synthetically lethal with a cold sensitive allele of *PRP16* (*prp16-301*) and not a temperature sensitive allele (*prp16-2*) or a branchsite suppressor allele (*prp16-101*). We have also tested a *prp16-1 prp8-101* double mutant and found a strong synthetic growth defect but not synthetic lethality (data not shown). *prp16-1* is similar to *prp16-101* in its ability to suppress mutant branchsites and to *prp16-301* in its cold sensitive, dominant negative phenotype. However, the phenotype of *prp16-1* is not as severe as that of *prp16-301*. As *prp16-101* suppresses branchsite mutations as well or better than *prp16-1* (Burgess and Guthrie, 1993), branchsite suppression *per se* is not the basis for the genetic interaction we see between *prp16-1* and *prp8-101*. Instead, the severity of the genetic interactions between *prp16-1* or *prp16-301* and *prp8-101* correlates with the severity of the cold sensitive, dominant negative phenotypes caused by the *prp16* alleles. These phenotypes appear to stem from the non-productive binding of mutant Prp16p to spliceosomes (Schwer and Guthrie, 1992b; Madhani and Guthrie, 1994b). *prp16 -1* inhibits the association of Prp8p with the 3' splice site and the same is likely to be true of *prp16-301*. The other genetic interactions between *prp8-101* and second step mutants might also be explained by the biochemical phenotypes we observe. Individually, each of these mutants impairs binding of Prp8p to the 3' splice site. When combined with *prp8-101* binding might be eliminated and, thus, cause a severe block to the second step of splicing *in vivo*.

Another explanation for the genetic interactions between *prp8-101* and the other second step splicing mutants is that the proteins they encode are physically associated in a complex that is only partly disrupted by individual mutant alleles but is rendered non-functional with two mutant subunits. Although immunodepletion and complementation

experiments argue against a stable extra-spliceosomal complex that contains these proteins (Schwer and Guthrie, 1991; Horowitz and Abelson, 1993a; Athar and Schwer, 1995; Jones et al., 1995), they may associate only on the spliceosome. As Prp8p is an integral and highly conserved spliceosomal protein (Anderson et al., 1989; Pinto and Steitz, 1989; Hodges et al., 1995), it is a good candidate for forming part of a spliceosomal binding site for the exchangeable second step factors. Whether or not they function as a complex, it is clear from the 3' splice site binding studies above that there is a great deal of functional interdependence among the second step splicing factors.

The 3' splice site is recognized in two distinct steps

While both Prp16p and Slu7p are predicted to be RNA binding proteins, Prp16p associates with spliceosomes only transiently and the same is likely to be true of Slu7p (Schwer and Guthrie, 1991; D. Frank, unpublished observations). Furthermore, their affinities for spliceosomal RNAs are probably highly regulated, and binding may require a very specific conformation of spliceosomal RNAs and proteins that would be difficult to reproduce with purified proteins and RNA. As an alternative strategy, site-specific UV crosslinking with a kinetically slowed substrate has allowed us to identify specific interactions between the 3' splice site and Prp16p and Slu7p in active spliceosomes.

Using splicing extracts that are blocked at specific steps of the reaction, we determined when Prp8p, Prp16p and Slu7p interact with the 3' splice site. With an extract blocked at the Prp2p-requiring step of splicing, we established that the 3' splice site interaction of these proteins takes place immediately prior to or, most likely, after the first catalytic step. This finding is important since it establishes a correlation between the time of 3' splice site interaction for Prp8p, Prp16p and Slu7p and the time in the splicing reaction when they are functionally required. This result has also been seen for Prp8p using site-specific 4-thio-uridine crosslinking at the 3' splice site (Teigelkamp et al., 1995a).

We have also shown that Prp16p interacts with the 3' splice site after the first catalytic step and prior to its hydrolysis of ATP. Although Prp8p and Slu7p can crosslink to the 3' splice site prior to Prp16p binding to the spliceosome, they require hydrolysis of ATP by Prp16p for maximal interaction with the 3' splice site. It is possible that the weaker 3' splice crosslinking of Prp8p and Slu7p in the absence of Prp16p and ATP is due to incomplete removal of Prp16p by immunodepletion. However, we also observe residual binding of Prp8p to the 3' splice site in heat-inactivated extracts prepared from a thermolabile mutant strain, *prp16-2* (data not shown). Because this experiment involves a completely independent means of removing Prp16p activity, the weaker 3' splice site interaction we observe for Prp8p and Slu7p in ΔPrp16 extracts is unlikely to be artifactual. This result is interesting in light of the observation that in *prp16-1*-derived extracts, in which the mutant Prp16p binds tightly to spliceosomes (Schwer and Guthrie, 1992b), crosslinking of Prp8p and Slu7p is largely precluded. It appears from this result that the Prp16p-3' splice site interaction is mutually exclusive with the 3' splice site interactions of Prp8p and Slu7p. In contrast, purified wild type Prp16p added to ΔPrp16 spliceosomes in the absence of ATP does not seem to affect the 3' splice site interaction of Prp8p and Slu7p. Therefore, wild type Prp16p must either be more exchangeable than the *prp16-1* encoded protein (Schwer and Guthrie, 1992b) or must bind in a conformation that does allow Prp8p and Slu7p to interact with the 3' splice site.

In summary, these results establish that the 3' splice site interaction of Prp16p and ATP hydrolysis by Prp16p precede a strong 3' splice site interaction with Prp8p and Slu7p. The weaker interaction with the 3' splice site by Prp8p and Slu7p in the absence of Prp16p may represent a 3' splice site proofreading or inspection step that is required prior to the conformational change induced when Prp16p hydrolyzes ATP. Interestingly, Prp16p, Slu7p and Prp8p can interact with the 3' splice site prior to ATP hydrolysis by Prp16p but do not confer protection to the 3' splice site from oligonucleotide directed RNase H cleavage (Schwer and Guthrie, 1992a). After ATP hydrolysis by Prp16p, either

additional factors must bind the 3' splice site and/or Prp8p and Slu7p must bind in a manner that allows increased protection.

The transition from weak to strong 3' splice site binding by Prp8p and Slu7p might correspond to the LI -> LI* transition proposed by Burgess and Guthrie (1993) as a key step in a kinetic pathway for maintaining fidelity of intron recognition. Both Prp8p and Slu7p are known to affect 3' splice site selection (Frank and Guthrie, 1992; Umen and Guthrie, 1995) and might need to "examine" a potential 3' splice site in a low affinity binding mode before a decision is made regarding utilization of that splice site. This low affinity binding might explain the apparent lack of sequence specificity seen in experiments where Prp8p can be crosslinked to 4-thio-uridine at non-utilized 3' splice sites (Teigelkamp et al., 1995a). Although crosslinking to a position that is formally +13 in the second exon was observed, this crosslink is in a region of the substrate that can be utilized as a 3' splice acceptor when there is no upstream 3' splice site. In contrast, we see no Prp8p crosslinking to positions downstream of +9 in the second exon of XL7 (Umen and Guthrie, 1995) and, correspondingly, this region of the substrate is never used as a 3' splice acceptor *in vitro* or *in vivo* (Umen and Guthrie, unpublished observations). It is also possible that the discrepancy we see is due to the exons of different splicing substrates having differing affinities for Prp8p.

Our results with *prp17* and *prp18* mutant extracts and ΔSlu7 immunodepleted extracts are consistent with Prp17p, Prp18p and Slu7p being required for strong interaction of Prp8p with the 3' splice site and with Prp17p and Prp18p being required for strong interaction of Slu7p with the 3' splice site. Furthermore, it appears that release of Prp16p from the 3' splice site is dependent on Slu7p and also possibly Prp18p. This requirement cannot be absolute for Prp18p since this protein is not essential for splicing *in vitro* or *in vivo* (Horowitz and Abelson, 1993a; Horowitz and Abelson, 1993b). Notably, both Slu7p and Prp18p act at an ATP-independent step of splicing (Horowitz and Abelson, 1993a; Athar and Schwer, 1995; Jones et al., 1995), which formally places their functions

"downstream" of the Prp16p-dependent step. However, it is clear that there exists an interdependent relationship between Slu7p and Prp16p: Prp16p requires Slu7p for 3' splice site release and Slu7p requires Prp16p to interact strongly with the 3' splice site. This mutual dependence might serve to couple the functions of the two proteins and enhance the fidelity or specificity of 3' splice site selection.

In summary, Prp16p crosslinking to the 3' splice site occurs after the first catalytic step of splicing and is independent of ATP binding or hydrolysis. Maximal 3' splice site crosslinking by Prp8p and Slu7p occurs after Prp16p hydrolyzes ATP and, presumably, after it exits the spliceosome. Thus, the 3' splice site is recognized in two distinct steps (see Figure 6). This result is consistent with experiments in mammalian extracts that also suggest at least two separate 3' splice site recognition events (Reed, 1989; Zhuang and Weiner, 1990). However, unlike yeast, many mammalian introns require the 3' splice site to initiate spliceosome assembly and at least one of the mammalian 3' splice site recognition events is likely to occur prior to the first catalytic step (Reed and Maniatis, 1985; Ruskin and Green, 1985; Rymond and Rosbash, 1985; Lamond et al., 1987; Rymond et al., 1987). In contrast, the 3' splice site recognition events that we have analyzed occur exclusively during the second catalytic step. Furthermore, the 3' splice site crosslinking interactions of Prp8p and Slu7p define at least one additional event that occurs after both the ATP dependent stage of the second catalytic step (defined by Prp16p and Prp17p), and the ATP independent stage (defined by Slu7p and Prp18p) but before catalysis. We cannot yet distinguish whether Prp8p and Slu7p crosslink to the 3' splice site sequentially or simultaneously. In either case, Prp8p and Slu7p bind the 3' splice site at the closest interval to catalysis that can presently be identified (Figure 6).

Although catalysis in pre-mRNA splicing is thought to be mediated primarily by snRNAs, participation by protein factors has not been ruled out. As Prp8p is, to date, the most evolutionarily conserved protein in the spliceosome (Anderson et al., 1989; Pinto and Steitz, 1989; Hodges et al., 1995), it is a good candidate for being involved in the catalytic

steps of the reaction. Our results strengthen this hypothesis by establishing the presence of Prp8p at or near the active site just prior to and/or during the second catalytic step of the reaction.

MATERIALS AND METHODS

Strains

Yeast strains are listed in Table 2. *E. coli* strain DH5α was used as a recipient for all cloning procedures described.

Genetic Methods

Standard yeast genetic methods were used for the manipulations described here (Guthrie and Fink, 1991). For double mutant analysis, plasmids pJU190 (*prp8-101*) or pJU213 (*prp8-1*) were cut with *HpaI* and transformed into recipient yeast strains carrying a *prp* mutation (either chromosomally or on a single copy plasmid complementing a chromosomal deletion) and a *URA3*-marked plasmid with a wild type copy of the mutant *PRP* gene. For pJU190 40-60% of His^+ transformants displayed the *prp8-101* phenotype (cold sensitivity, altered 3' splice site usage) and for pJU213, 80-90% of His^+ transformants displayed a *prp8-1* phenotype (temperature sensitivity). Approximately 60 individual transformants were examined for each experiment. These transformants were replica plated to 5FOA-containing plates and scored for viability. As the integration frequency of *prp8-101* with pJU190 is ~50%, this frequency of $5FOA^-$ tranformants is indicative of synthetic lethality. For pJU190 either ~50% of the transformants did not grow on 5FOA, indicating synthetic lethality with the *prp* mutant being tested, or 98-100% grew on 5FOA, indicating no genetic interaction. For pJU213, 98-100% of the transformants always grew on 5FOA indicating no genetic interactions with *prp* mutants that were tested. Double mutants were assayed at the permissive temperature of 25°.

Plasmids

Plasmid pJU204 (*PRP8-HA3*) has been described (Umen and Guthrie, 1995). pJU206 (*prp8-101-HA3*) was constructed from pJU204 by replacing the 1.9 kb *ClaI* fragment from pJU204 with the same fragment from a *prp8-101* clone. The *prp8-101*

integrating plasmid pJU190 was constructed by inserting a SalI-NheI fragment from a YCP50 *prp8-101* clone into the *SalI* and *XbaI* sites of the *HIS3* marked integrating plasmid RS303 (Sikorski and Hieter, 1989). The *prp8-1* version of this plasmid was constructed by PCR amplifying the 1.9 kb C-terminal *ClaI* fragment of *PRP8* from a *prp8-1* strain and then replacing the 1.9kb*ClaI* fragment of pJU190 with the PCR amplified *prp8-1* fragment. Plasmids for double mutant analysis were from the following sources: *PRP2* (Chen and Lin, 1990), *PRP3* (Last et al., 1987), *PRP24* (Shannon and Guthrie, 1991), *prp24-6* (gift from Anita Jandrositz, Univ. of Calif., San Francisco), *PRP17* (Jones et al., 1995), *PRP18* (Horowitz and Abelson, 1993b), *SLU7*, and *slu7-ccss*(Frank and Guthrie, 1992) *PRP16*, *prp16-2*, *prp16-301* and *prp16-101*(Burgess and Guthrie, 1993)

Immunodepletions

Prp16p was removed from a wild type extract containing epitope-tagged Prp8p (PRP8-HA3) by incubating 90 µl of extract with 18 µl (288 µgms) protein A-purified αPrp16p antibodies (Schwer and Guthrie, 1991) at 4° for 1 hour. This reaction was added to 120 µl protein A Sepharose beads in buffer D plus an additional 75 µl buffer D (Lin et al., 1985), and incubated at 4° for an additional hour with frequent agitation. After brief centrifugation, the supernatant was removed and used for splicing reactions. Slu7p was removed from extracts by a similar procedure except 80 µl of αSlu7p antibodies (960 µgms) (Jones et al., 1995) were used and no additional buffer D was added.

Splicing, UV Crosslinking and Immunoprecipitations

Splicing extracts were prepared as described previously (Umen and Guthrie, 1995). For mutant *prp* strains, an epitope-tagged version of *PRP8* (pJU204) was transformed into the strains prior to growth and extract preparation. Cells were grown in SD -His medium to maintain selection for the plasmid until the last 2-3 generations. At this point they were transferred to YEPD where they were grown until harvesting.

Splicing reactions were performed as previously described (Umen and Guthrie, 1995). When mutant extracts were complemented, they were first incubated for 20 minutes at 25°. A 300- to 500-fold molar excess of cold competitor RNA was then added with either a half reaction volume of wild type complementing extract plus standard splicing salts or additional mutant extract plus standard splicing salts. The incubation was then continued for 10 minutes. For each experiment, an aliquot of the complementing extract mixture was removed before splicing commenced and mixed with cold competitor RNA as a control. The *prp2-1* derived mutant extract was used with or without prior heat inactivation for 10 minutes at 37°. ΔPrp16 extracts were incubated for 20 minutes and then incubated at 30° for 10 minutes in the presence of 4 mM glucose and 8.5 units/ml of hexokinase (Boehringer). The reaction was then split into four parts which received either buffers only, 2.5 mM additional ATP and MgCl2, 2.5 pg/μl purified Prp16p in Buffer D, or both ATP/MgCl2 and Prp16p. The incubation was then continued for 10 minutes at 25°.

UV crosslinking and immunoprecipitations were performed as described previously (Umen and Guthrie, 1995) except 150 mM NaCl (instead of 300 mM KCl) was used in the wash buffers and no salt or NP40 were added to splicing reactions during incubation with antiserum. 12 μgm αSlu7p antibodies or 8μgm αPrp16p antibodies (protein A purified) were used for immunoprecipitations from 40 μl reactions. αPrp17p (50 μl), αPrp18p (60 μl) and αSnp1p (80 μl) serum were pre-bound to protein A Sepharose (40 μl)(Pharmacia) and incubated with crosslinking reactions (40 μl) in 500μl NET150 (150 mM NaCl, 50 mM Tris 7.4, .05% NP40) for two hours at 4° prior to elution and SDS-PAGE. After electrophoresis, the gels were transferred to nitrocellulose and autoradiographed or scanned using a phosphorimager. In some experiments the nitrocellulose was then immunoblotted.

For denaturing immunoprecipitations, a splicing reaction containing epitope-tagged wild type Prp8p was divided in half. One half was subject to UV crosslinking and the other was not. SDS was added to a final concentration of 5% and the samples were boiled for 3 minutes. Each sample was diluted 10-fold into NET150 that contained 1% Triton X-

100, 1 μg 12CA5 antibody, 15 μl protein A Sepharose and protease inhibitors. After 1 hour rocking at 4°, the beads were washed 3 times with NET150 and then treated with proteinase K and SDS. RNA was precipitated after phenol/chloroform extraction and fractionated on a 6% denaturing polyacrylamide gel.

snRNA immunoprecipitation controls (Figure 3E) for Prp8p, Prp18p, and Snp1p involved mock splicing reactions that were identical to normal reactions but did not contain XL7 or undergo UV crosslinking and RNase treatment. The pellets from the immunoprecipitate were proteinase K treated and extracted and the RNA was precipitated and probed using oligos directed against U1, U4, U5 and U6 snRNAs (Bordonne et al., 1990).

The immunoprecipitation control for Prp17p was performed by TCA precipitating the supernatant from a crosslinking/immunoprecipitation experiment using 10% TCA. This was resuspended in protein sample buffer and compared to an equivalent amount of total starting extract before immunoprecipitation. These samples were fractionated by SDS-PAGE and immunoblotted to detect Prp17p. Some antibodies remain in the supernatant after immunoprecipitation accounting for the signal beneath the Prp17p band on the gel in figure 3D.

All immunoblotting was performed using an ECL kit from Amersham according to the manufacturer's instructions.

ACKNOWLEDGMENTS

We thank Liz Blackburn, Cathy Collins, Amy Kistler, Hiten Madhani, Chris Siebel, Peter Walter and Yan Wang for comments on this manuscript. We thank Yan Wang for providing purified Prp16p and R. J. Lin for providing purified Prp2p. We thank Shelly Jones for providing αPrp17p serum and recombinant Prp17p, and Dan Frank for providing αSlu7p serum. We thank Paul Siliciano for providing aSnp1p serum and David Horowitz for providing αPrp18p serum. We are grateful for the excellent technical assistance of Heli Roiha, Carol Pudlow and Lucita Esperas.

This work was supported by a NIH training grant and a NSF pre-doctoral fellowship to J.G.U., and NIH research grant GM21119 to C.G. C.G. is an American Cancer Society Research Professor of Molecular Genetics.

Figure 1. Strategy for analyzing 3' splice site binding proteins.

(A) Schematic of XL7 *in vitro* splicing substrate. The 3' splice site nucleotides are shown with the A to G point mutation and a single labeled phosphate depicted by a * symbol. The 15 nucleotide RNase T1 digestion product is underlined. (B) XL7 is incubated in splicing extract and bound proteins are UV crosslinked. After digestion with RNase T1, only 3' splice site bound proteins will be labeled and specific proteins can be isolated by immunoprecipitation.

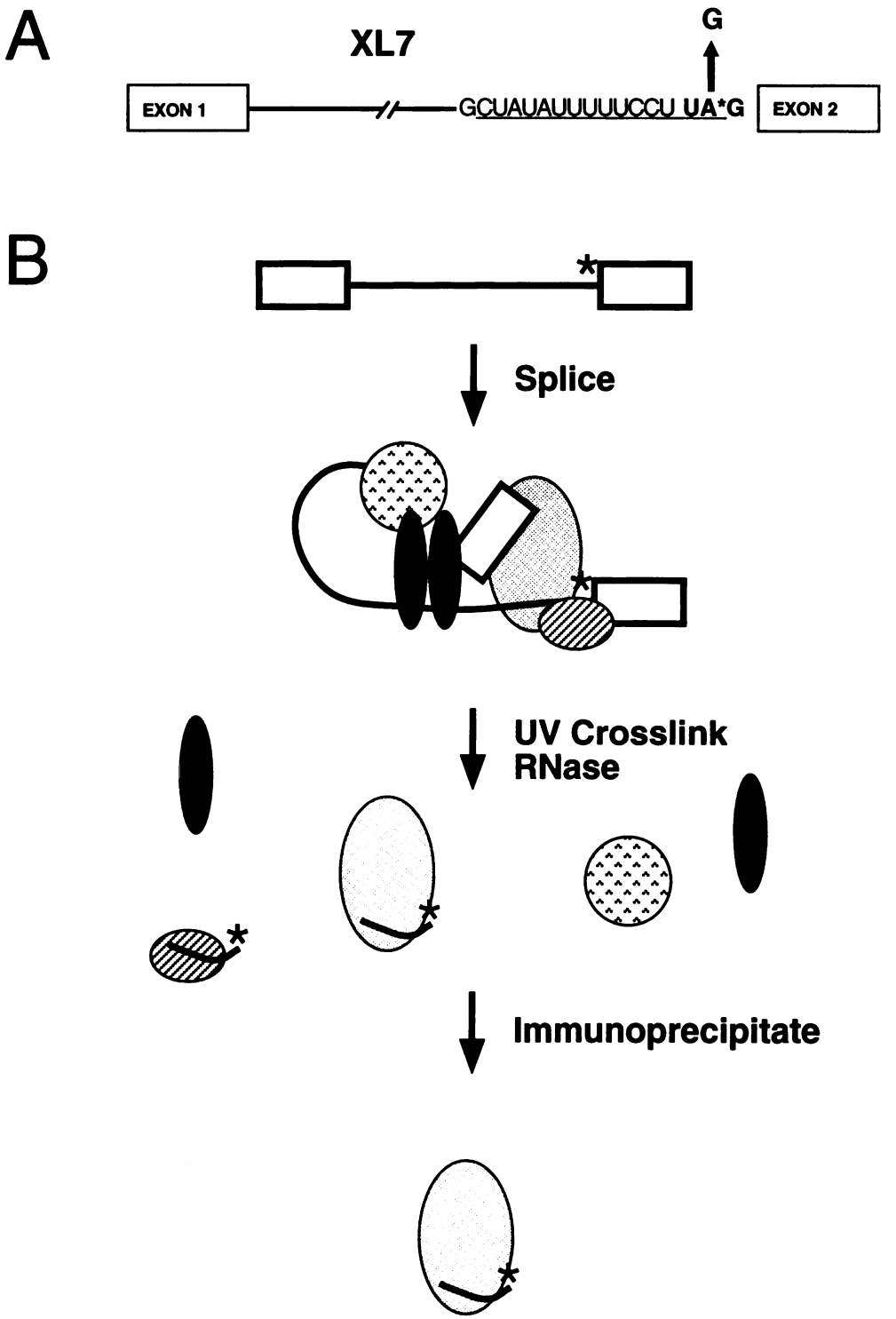

A
XL7
G
EXON 1
GCUAUAUUUUUCCU UA*G
EXON 2
B
Splice
UV Crosslink
RNase
Immunoprecipitate

Figure 2. Analysis of 3' splice site crosslinking in *prp8-101* strains.

(A) Splicing of XL7 in epitope-tagged *PRP8* (lane 1) and epitope-tagged *prp8-101* (lane 2) extracts. *HpaII* digested pBR322 markers are in lane m. Products of splicing are cartooned to the left. (B) Phosphorimage analysis of the crosslinked and immunoprecipitated Prp8p from (A). (C) Immunoblot analysis of the epitope-tagged Prp8p from (B). (D) Native immunoprecipitation of spliceosomes from (A) with anti-HA antibodies (lanes 2 and 3) and a wild type, non-epitope- tagged extract in lane 1. (E) Denaturing immunoprecipitation of wild type, epitope-tagged Prp8p from reaction in (A) with (lane 1) and without (lane 2) UV crosslinking.

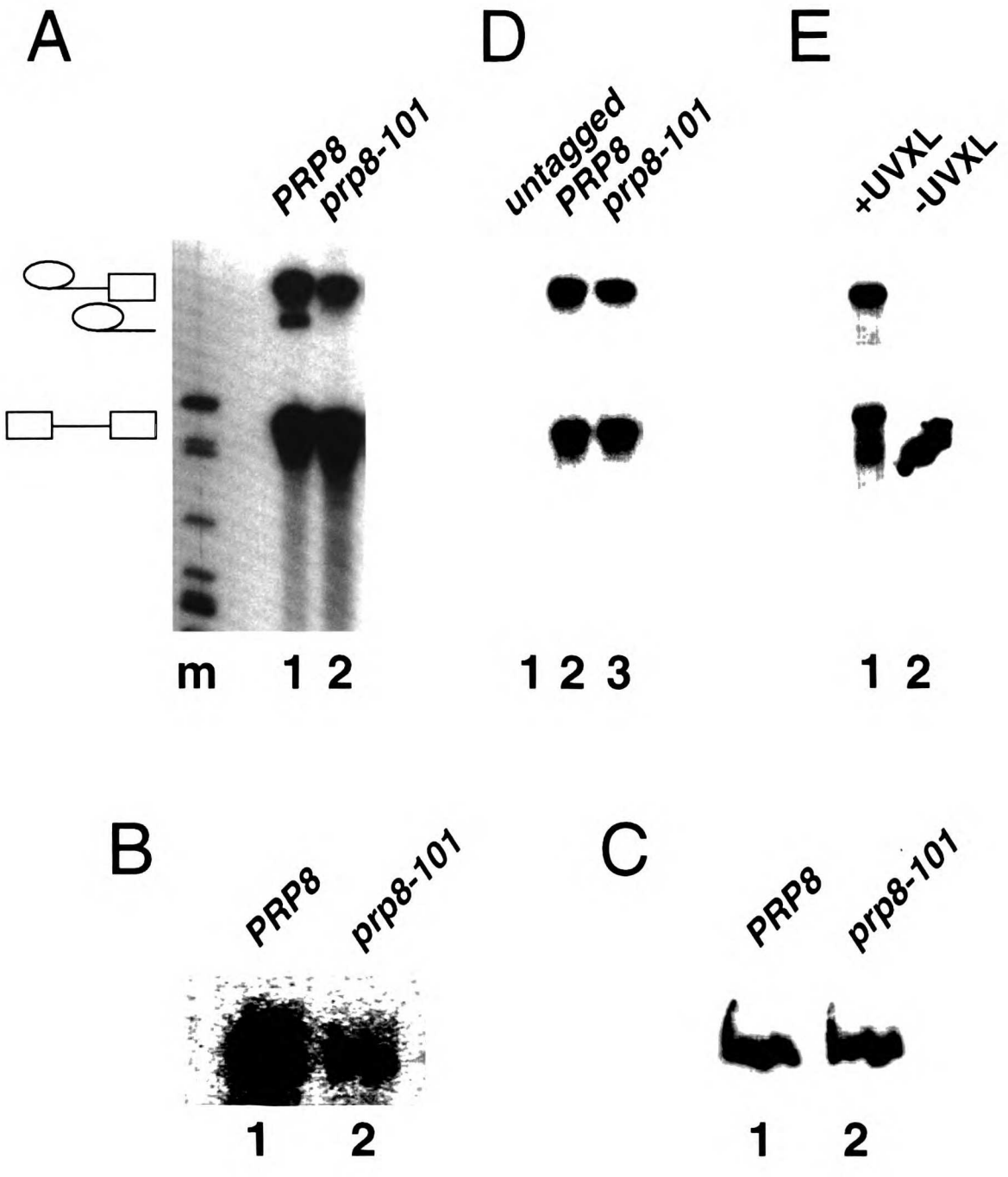
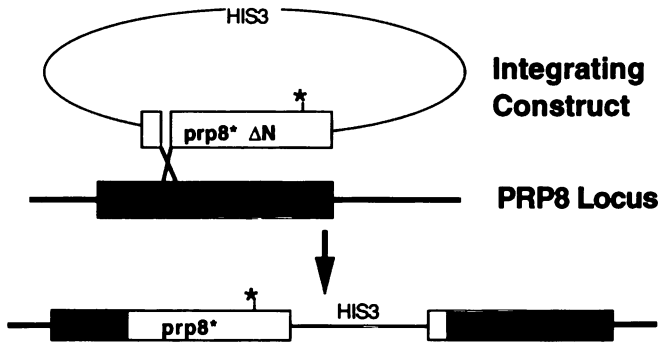
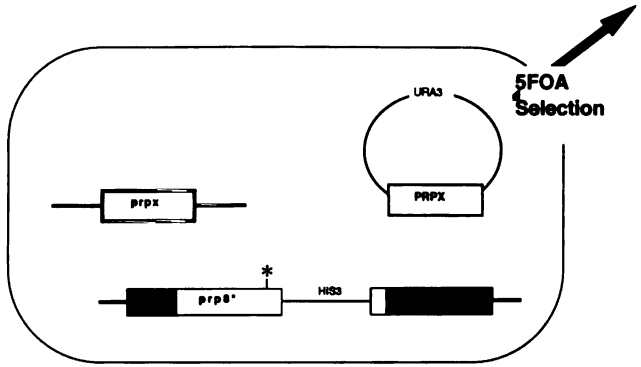

A
PRP8
prp8-101
m 1 2
D
untagged PRP8
prp8-101
1 2 3
E
+UVXL
-UVXL
1 2
B
PRP8
prp8-101
1 2
C
PRP8
prp8-101
1 2

Figure 3. Genetic interactions between *prp8-101* and second step mutants. (A) A gapped, integrating, *HIS3*-marked plasmid containing an N-terminal truncation of *prp8-1* or *prp8-101* (open box) is shown recombining with the chromosomal *PRP8* locus (shaded box). The mutation is indicated by a * symbol. The integration produces one full length copy of *prp8-1* or *prp8-101* and a truncated, non-functional copy of *PRP8, PRP8ΔN*. (B) The *prp8* mutants from (A) were introduced directly into cells containing a second *prp* mutation (*prpx*) that is complemented by a wild type (*PRPX*), *URA3*-marked plasmid. Selection on 5FOA generates a *prp8-1 prpx* or *prp8-101 prpx* double mutant strain (See MATERIALS AND METHODS for details). (C) (+) symbol indicates that the double mutant combination is viable and (-) indicates that the double mutant combination is inviable and that the two mutations are synthetically lethal.

A



B



C

Viability of Double Mutants

	prp2-1	*prp3-1*	*prp24-6*	*prp16-101*	*prp16-2*	*prp16-301*	*prp17-1*	*prp17-2 (slu4-1)*	*prp18-1*	*slu7-1*	*slu7-ccss*
prp8-1	+	+	+	+	+	+	+	+	+	+	+
prp8-101	+	+	+	+	+	-	-	-	-	-	-

Figure 4. Crosslinking of Prp16p and Slu7p to the 3' splice site. (A) Splicing of XL7 (lanes 1 and 4) and derivatives (see text). Products of splicing are cartooned. The single phosphate label in XL7 allows detection of only precursor, lariat intermediate and excised lariat intron. (B) A splicing reaction (panel A, lane 1) was crosslinked, and proteins were immunoprecipitated with antiserum (indicated above each lane) and fractionated by SDS-PAGE and autoradiographed. The bands corresponding to Prp8p, Prp16p and Slu7p are indicated by a filled circle. The presumptive locations of Prp17p, Snp1p and Prp18p are indicated by open circles. The crosslinked proteins at ~66 kD, ~150 kD and ~220 kD (most prominent in lanes 4-6) are non-specific and variable contaminants in immunoprecipitations. Positions of molecular weight markers are on the right. (C) Immunoprecipitation reactions similar to those in (B) except without crosslinking or RNase T1 digestion were protease treated and extracted. The RNAs were fractionated and probed for U1, U4, U5 and U6. Antiserum used in lanes 1-3 are indicated above each lane and compared to a mock reaction (no antiserum) in lane 4. The positions of the snRNAs are indicated. The bands not corresponding to these are non-specific contaminants. Some U6 snRNA was immunoprecipitated non-specifically in this experiment. (D) Immunoprecipitation of Prp17p was evaluated by comparing the input from a splicing reaction as in (B) (lane 1) to the supernatant from the immunoprecipitation (lane 2). Some IgG remains in the supernatant. Recombinant Prp17p (lane 3) was used as a size standard. (E) Samples from (A) were crosslinked and immunoprecipitated with antiserum against Slu7p or Prp16p, fractionated by SDS-PAGE, transferred to nitrocellulose and phosphorimaged. Only the region of the gel containing Prp16p or Slu7p is shown. (F) The samples in (E) lanes 4 and 5 were subsequently immunoblotted with the appropriate serum indicated above the lanes. Positions of Prp16p, Slu7p and IgG are indicated. Positions of molecular weight markers are on the left.

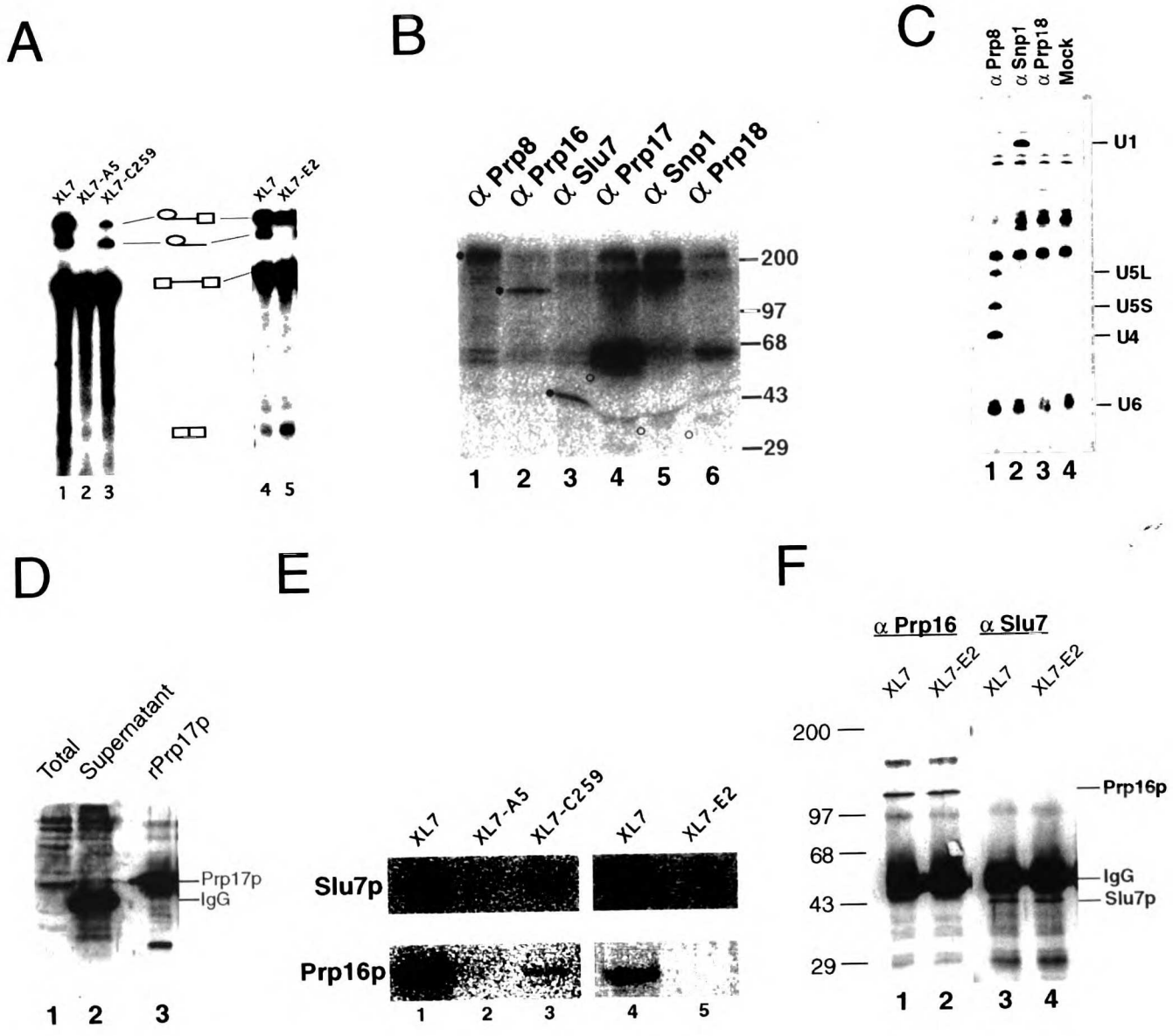
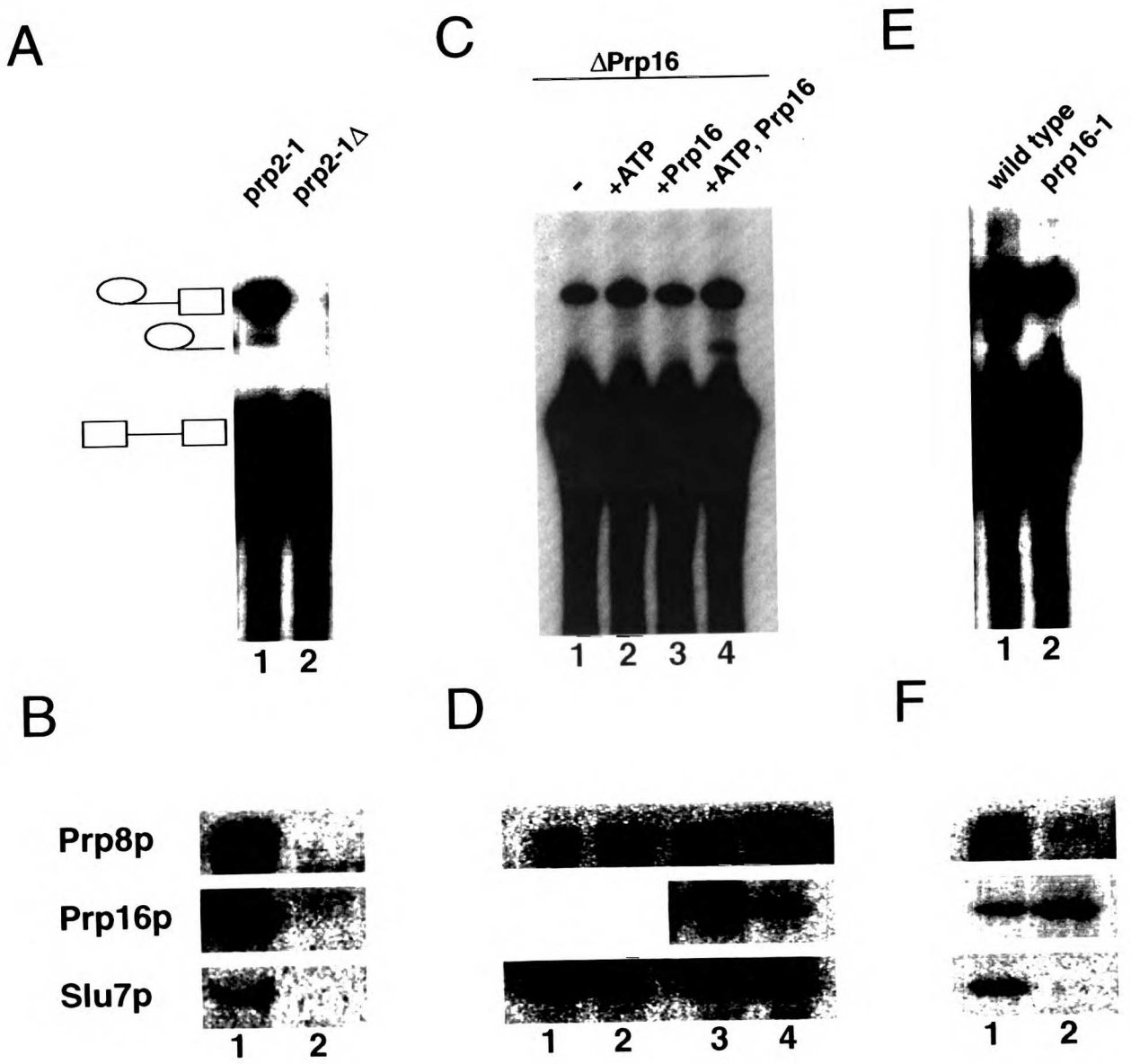

A
XL7
XL7-A5
XL7-C259
XL7
XL7-E2
1 2 3
4 5
B
α Prp8
α Prp16
α Slu7
α Prp17
α Snp1
α Prp18
200
97
68
43
29
1 2 3 4 5 6
C
α Prp8
α Snp1
α Prp18
Mock
U1
U5L
U5S
U4
U6
1 2 3 4
D
Total
Supernatant
rPrp17p
Prp17p
IgG
1 2 3
E
XL7
XL7-A5
XL7-C259
XL7
XL7-E2
Slu7p
Prp16p
1 2 3 4 5
F
α Prp16
α Slu7
XL7
XL7-E2
XL7
XL7-E2
200
97
68
43
29
Prp16p
IgG
Slu7p
1 2 3 4

Figure 5. (A) Splicing of XL7 in a *prp2-1* extract without (lane 1) and with (lane 2) prior heat inactivation. Products of splicing are diagrammed to the left of the gel. (B) samples from the reactions in (A) were crosslinked and immunoprecipitated with serum against epitope- tagged Prp8p or against Prp16p or Slu7p and scanned with a phosphorimager. Only the region of the gel where the protein of interest migrates is shown. (C) Splicing in Prp16p and ATP depleted extracts with no additions (lane 1), ATP added (lane 2), Prp16p added (lane 3) or ATP and Prp16p added (lane 4). (D) Crosslinking of samples from (C) analyzed as in (B). (E) Splicing of XL7 in a wild type (lane 1) or *prp16-1* extract (lane 2). (F) Crosslinking of samples from (E) analyzed as in (B). (G, I, K) Splicing in staged reactions with *prp17-1*, *prp-18-1* derived or Slu7p immunodepleted (ΔSlu7) extracts. Splicing was initiated in mutant or depleted extracts without cold competitor RNA (First Addition, lanes 1 and 3) to allow the first step of splicing. Additional mutant or depleted extract (Second Addition, lane 1) or complementing wild type extract (Second Addition, lane 3) was added along with excess cold competitor RNA prior to a second incubation. Lane 2 contains cold competitor RNA added at the beginning of a similar reaction as in lanes 1 and 3. Products of splicing are cartooned to the left of (G). (H, J, L) Samples from (G, I, K; lanes 1 and 3) were analyzed as in (B) for crosslinking to Prp8p, Prp16p and Slu7p after the second incubation period. Lane 1 represents the uncomplemented reaction and lane 2 represents the complemented reaction.


A
prp2-1
prp2-1Δ
1 2
C
ΔPrp16
-
+ATP
+Prp16
+ATP, Prp16
1 2 3 4
E
wild type
prp16-1
1 2
B
Prp8p
Prp16p
Slu7p
1 2
D
1 2 3 4
F
1 2

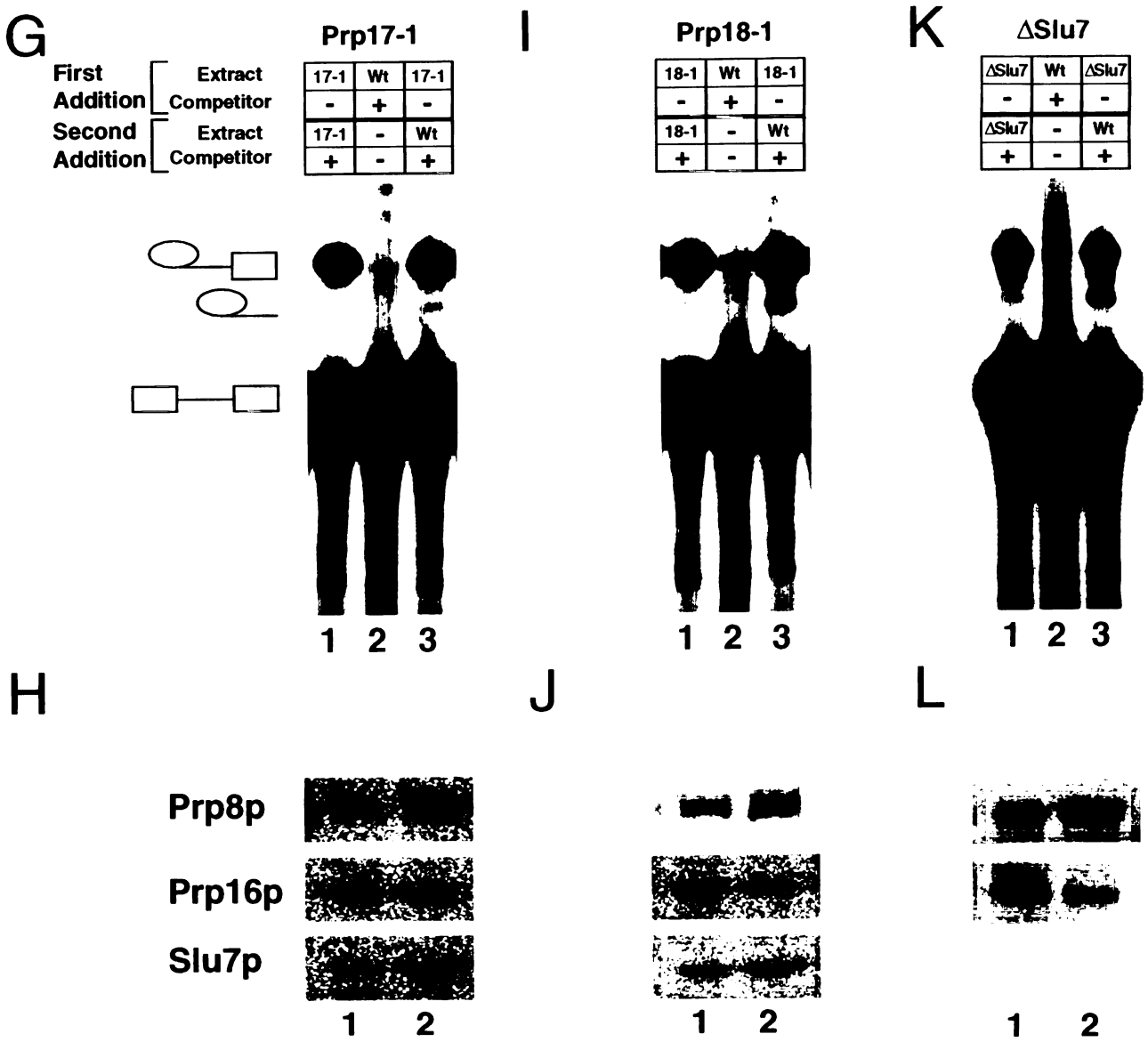

G
Prp17-1
First Addition
Extract
Competitor
Second Addition
Extract
Competitor
17-1 Wt 17-1
- + -
17-1 - Wt
+ - +
1 2 3
I
Prp18-1
18-1 Wt 18-1
- + -
18-1 - Wt
+ - +
1 2 3
K
ΔSlu7
ΔSlu7 Wt ΔSlu7
- + -
ΔSlu7 - Wt
+ - +
1 2 3
H
Prp8p
Prp16p
Slu7p
1 2
J
1 2
L
1 2

Figure 6. Model for ordered interactions of Prp8p, Prp16p and Slu7p with the 3' splice site. A summary of a splicing reaction starting with the Prp2p dependent step and indicating gene product and ATP hydrolysis requirements. Proteins bound to the 3' splice site are indicated by the spheres labeled 16 (Prp16p), 7 (Slu7p) and 8 (Prp8p). After the Prp2p step, Prp16p binds strongly to the 3' splice site and Slu7p and Prp8p can bind weakly. After hydrolysis of ATP by Prp16p, and the Prp17p-, Prp18p-, and Slu7p-dependent steps, Prp8p and Slu7p bind maximally to the 3' splice site. Prp17p is shown acting at the Prp16p-dependent step since Prp17p is known to function at or before an ATP requiring step (Jones et al., 1995) but is not necessary for Prp16p to bind the 3' splice site. Slu7p and Prp18p function in a ATP-independent manner and are placed after the Prp16p- and Prp17p-requiring steps (Horowitz and Abelson, 1993a; Athar and Schwer, 1995; Jones et al., 1995). Prp8p is at the 3' splice site just prior to (and possibly during) catalysis and leaves the 3' splice site afterwards. The same is presumed to be true of Slu7p.

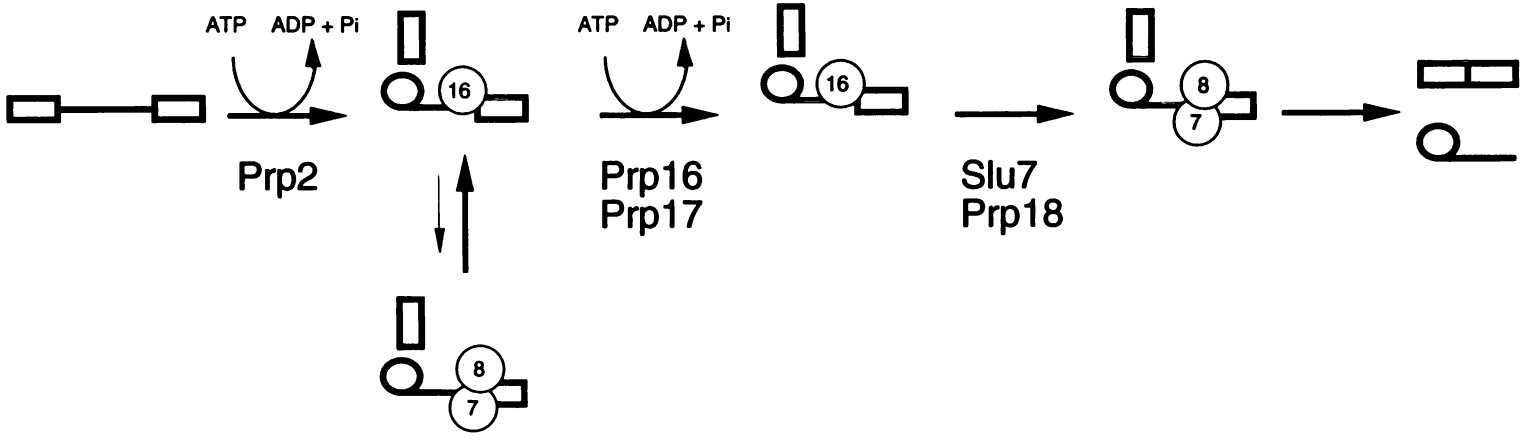

ATP
ADP + Pi
16
Prp2
ATP
ADP + Pi
16
Prp16
Prp17
Slu7
Prp18
8
7
8
7

Table 1. 3' splice site crosslinking to Prp8p, Prp16p and Slu7p in mutant or depleted extracts is expressed in comparison to a wild type or control (complemented) extract. (+) wild type levels of crosslinking. (+/-) less than wild type levels of crosslinking. (-) no detectable crosslinking. (++) greater than wild type levels of crosslinking. (n.d.) experiment not done. (n.a.) not applicable.

Table 1

Summary of 3' Splice Site Crosslinking to Prp8p, Prp16p and Slu7p

Extract	Prp8 x-link	Prp16 x-link	Slu7 x-link
Wild Type	+	+	+
prp8-101	+/-	n.d.	n.d.
prp2-1	-	-	-
ΔPrp16	+/-	n.a.	+/-
ΔPrp16 + ATP	+/-	n.a.	+/-
ΔPrp16 + 16	+/-	++	+/-
ΔPrp16 + ATP + 16	+	+	+
prp16-1	-	++	-
Prp17-1	+/-	+	+/-
Prp18-1	+/-	++	+/-
ΔSlu7	+/-	++	n.a.

Table 2

S. cerevisiae strains used in this study

Name	Genotype	Reference
YJU76	*MATα prp8Δ::LEU2 leu2-3 leu2-112 ura3-52 ade2-101 his3Δ1 trp1-289 prp4Δ::TRP1* pJU186 *(PRP8-HIS3-CEN-ARS)*	(Umen and Guthrie, 1995)
YJU77	YJU76 except contains plasmid pJU204 (PRP8-HA3) instead of pJU186	(Umen and Guthrie, 1995)
YJU78	YJU76 except contains plasmid pJU206 (prp8-101-HA3) instead of pJU186	This work
SS304	*MATα prp2-1 ade2-1 his3-532 trp1-289 ura3-1 ura3-2*	(Lustig et al., 1986)
TR1-3	*MATα his3 trp1 lys2-801 ura3-52 ade2-101 prp24Δ::LYS2* pYCpXba *(PRP24-URA3-CEN-ARS)*	Gift from Anita Jandrositz, Univ. of Calif., San Francisco
SPJ3.33	*MATa prp3-1 his3 leu2 lys2 ura3-52*	(Lustig et al., 1986)
YDAF7-GK	*MATa ura3 lys2 his3 ade2 trp1 leu2 slu7Δ::TRP1* pYS7-7 *(SLU7-URA3-CEN-ARS)*	(Frank and Guthrie, 1992)
A7C8Aa	*MATa slu7-1 ura3-52 trp1-Δ63 his3-Δ200 leu2-Δ1 ade2-101 lys2-801*	(Frank and Guthrie, 1992)
ts365	*MATa prp17-1 ade2-101 his3-Δ200 ura3-52 lys2-801*	(Vijayraghavan et al., 1989)
A4C15Ba	*MATa prp17-2 ura3-52 trp1-Δ63 his3-Δ200 leu2-Δ1 ade2-101 lys2-801*	(Frank et al., 1992)

ts503	*MATα prp18-1 ade2-101 his3-Δ200 ura3-52 lys2-801*	(Vijayraghavan et al., 1989)
YS78	*MATa trp1 ura3 lys2 leu2 ade2 his3 prp16Δ::LYS2* pSB2 *(PRP16-URA3-CEN-ARS)*	(Burgess and Guthrie, 1993)

CHAPTER 3

Mutagenesis of the Yeast Gene *PRP8* Reveals Domains Governing the Specificity and Fidelity of 3' Splice Site Selection

ABSTRACT

PRP8 encodes a highly conserved U5 snRNP protein required for spliceosome assembly prior to the first catalytic step of pre-mRNA splicing. We recently identified a novel allele, *prp8-101*, that defines a role for the encoded protein in 3' splice site selection and the second catalytic step. This mutant specifically impairs recognition of the uridine tract that precedes most yeast 3' splice sites and exacerbates the effects of point mutations at the 3' splice junction PyAG motif. In order to better define the domain(s) of *PRP8* that is responsible for uridine tract recognition, we carried out extensive mutagenesis of the gene and selected for new alleles that confer a phenotype similar to that of *prp8-101*. The 11 strongest alleles (including new isolates of *prp8-101*) cause changes in one of two amino acids in the C-terminal portion of the protein. We also identified a second class of *PRP8* mutant that affects the fidelity of 3' splice site utilization. In contrast to *prp8-101*, these alleles suppress point mutations in the PyAG motif at the 3' splice site and do not alter uridine tract recognition. The strongest of these alleles map to a region directly upstream of the *prp8-101*-like mutations. These new *PRP8* alleles define two separable functions of Prp8p, required for specificity of 3' splice site selection and fidelity of 3' splice site utilization, respectively. Taken together with other recent biochemical and genetic data, our results suggest that Prp8p plays a functional role at the active site of the spliceosome during the second catalytic step of splicing.

INTRODUCTION

Nuclear pre-mRNA splicing involves the recognition and removal of introns from messenger RNA precursors. Introns are identified by conserved sequences at the 5' splice site, branchsite, and 3' splice site. These sequences are recognized by five small nuclear ribonucleoprotein particles (U1, U2, U4, U5, and U6 snRNPs) which, together with numerous accessory proteins, assemble onto intron-containing RNAs to form the spliceosome. The spliceosome catalyzes the removal of introns in two chemical steps involving 5' splice site cleavage and branched lariat formation (step 1), followed by 3' splice site cleavage and exon ligation (step 2) (reviewed in Green, 1991; Guthrie, 1991; Rymond and Rosbash, 1992; Moore et al., 1993).

A key question is how introns are accurately identified. The 5' splice site and branchsite are recognized by a well-characterized set of interacting snRNAs and proteins (reviewed above). However, much less is known about how the 3' splice site is identified. In most organisms the 3' splice site is composed of a nearly invariant PyAG motif preceding the 3' splice junction and an upstream pyrimidine-rich tract. In mammals, the pyrimidine tract is first bound by U2AF, which is required for the first step of splicing, and later by PSF, which is required for the second step (Zamore & Green, 1991; Patton et al., 1993; Gozani et al., 1994).

Several RNA-RNA interactions are also required for proper 3' splice site selection. The first and last guanosine residues in introns share a non-Watson-Crick interaction that is critical for 3' splice site utilization (Parker & Siliciano, 1993; Chanfreau et al., 1994; Deirdre et al., 1995). U5 snRNA can interact with the first two residues of the second exon and may play a role in aligning the two exons during the second step of splicing (Newman & Norman, 1992; Sontheimer & Steitz, 1993). Because exon sequences are poorly conserved, this interaction probably does not normally play a major role in 3' splice site selection. Finally, mutations in U2 and U6 snRNAs can compromise the fidelity of 3' splice site utilization and nonspecifically suppress the effects of point mutations in the

PyAG motif (Lesser & Guthrie, 1993b; Madhani & Guthrie, 1994). As these residues are in a domain of U2 and U6 that is thought to be part of the spliceosomal active site, their alteration may change the architecture of this region so as to relax its stringency for 3' splice site nucleotide identity (Madhani & Guthrie, 1994).

We have focused on the role of a highly conserved U5 snRNP protein, Prp8p, in 3' splice site selection (Hodges et al., 1995; Umen & Guthrie, 1995a). Prp8p was first shown to have a role in spliceosome assembly prior to the first catalytic step (Jackson et al., 1988; Brown & Beggs, 1992). Recently, however, we identified a novel allele, *prp8-101,* that impairs recognition of the 3' splice site uridine-rich tract during the second catalytic step (Umen & Guthrie, 1995a). Consistent with a direct role in 3' splice site recognition, Prp8p can be crosslinked to the 3' splice site in a site-specific manner during splicing after the first catalytic step (Teigelkamp et al., 1995; Umen & Guthrie, 1995a). A more detailed kinetic analysis of its interaction with the 3' splice site indicates that Prp8p is likely to be bound to the 3' splice site during the second catalytic event (Umen & Guthrie, 1995b). However, to date, there is no functional evidence of a role for Prp8p in catalysis or in recognition of the conserved PyAG trinucleotide at the 3' splice junction.

Although Prp8p is likely to be an RNA-binding protein, it does not contain any significant homologies to known RNA-binding proteins or other families of proteins (Hodges et al., 1995). Thus, there are no obvious structural domains that might be candidates for intron binding sites. In this work, we have carried out an extensive mutagenesis of the *PRP8* gene with two objectives. The first is to genetically map the domain(s) of the protein responsible for uridine tract recognition, and the second is to determine whether a domain of Prp8p interacts functionally with the PyAG trinucleotide at the 3' splice site. We have found that all uridine recognition mutants (like *prp8-101*; Umen and Guthrie, 1995a) cause alterations in the C-terminal portion of the protein. The strongest of these new alleles (including new isolates of *prp8-101*) change one of two codons. We have identified a second class of mutations which alters the fidelity of 3'

splice site utilization. These also cause alterations in the C-terminal half of Prp8, but in a region upstream of the alterations caused by the *prp8-101*-like class. Unlike *prp8-101*, which exacerbates the effects of point mutations in the PyAG motif at the 3' splice junction (Umen & Guthrie, 1995a), this class of *PRP8* mutant suppresses the effects of PyAG alterations. The complex spectrum of preferences for different PyAG alterations displayed by these *PRP8* alleles is suggestive of a direct interaction between Prp8p and the PyAG trinucleotide and/or the spliceosomal active site.

MATERIALS AND METHODS

Yeast Methods: All methods for manipulation of yeast, including media preparation, growth conditions, transformation, plasmid recovery and 5FOA selection were performed according to standard methods (Guthrie & Fink, 1991). Copper growth assays and β-galactosidase assays were performed as previously described (Miller, 1972; Lesser & Guthrie, 1993a). Strains for β-galactosidase assays were grown in media containing 2% galactose and 2% raffinose for 24 hours prior to analysis in order to induce expression of the lacZ fusion construct. Strain YJU75 (with various plasmids described below) was used for all experiments: **a** *ade 2 cup1Δ::ura3 his3 leu2 lys2 prp8Δ::LYS2 trp1* pJU169 (*PRP8 URA3 CEN ARS*). Disruption of the *PRP8* locus is described below.

After mutagenesis and transformation, 3' splice site selection mutants were selected by replica plating to copper-containing plates before or after replica plating to 5FOA-containing plates.

Plasmid Construction: Molecular cloning procedures were carried out according to standard methods (Maniatis et al., 1982). Plasmid pJU225 (*PRP8 TRP1 2μ*) was constructed from a previously described plasmid JDY13 (*GAL::PRP8*) (Brown & Beggs, 1992). The galactose-driven promoter in this plasmid was removed by cutting with NheI and XhoI. A wild type copy of the promoter was amplified by PCR under non-mutagenic conditions and used to replace the galactose driven promoter. The entire *PRP8* gene was then excised using *XhoI* and *NotI*, and ligated to plasmid RS424 (*TRP1 2μ*) also cut with *XhoI* and *NotI* (Sikorski & Hieter, 1989).

For all non-mutagenic PCR reactions, Hot Tub polymerase (Amersham) was used according to the manufacturer's instructions.

The 2.1 kb *BstEII* fragment from pJU225 was swapped between wild type and mutant clones using standard procedures. The chimeras were confirmed by sequencing the relevant regions.

ACT1-CUP1 reporters are depicted in Figure 2. The 3' splice site competition reporters and construct set II were made according to previously described methods (Umen & Guthrie, 1995a). Construct set I plasmids were made by using oligonucleotide directed mutagenesis. The sequences that were altered compared to the standard *ACT1-CUP1* fusion are depicted.

The *ACT1-CUP1* G5A reporter with either the normal (Nl) or cryptic (Ab) 5' cleavage sites in frame with the *CUP1* coding sequence were constructed by Amy Kistler. The *rp51a-lacZ* fusion construct and *ACT1-CUP1* A259C construct have been described previously (Lesser & Guthrie, 1993a; Chanfreau et al., 1994)..

The *prp8Δ::LYS2* disruption plasmid was constructed by first eliminating the *Not I* site in the polylinker region of pJU225. A *NotI* site was then introduced into the 5' end of *PRP8* at the fifth codon using PCR-based mutagenesis. A 4 kb fragment containing the *LYS2* gene in the polylinker of a Bluescript plasmid (Stratagene) was excised with *NotI* and *ClaI* and ligated to the *ClaI* and *NotI* sites of the modified *PRP8* plasmid just described. This disruption removes all but the first five and last four codons of *PRP8*.

To generate strain YJU75, the *prp8Δ::LYS2* fragment with *PRP8* flanking sequence (1-2 kb on each side) was excised from the disruption plasmid with *SacI* and *ApaI* and transformed into strain L5 (Umen & Guthrie, 1995a), which contained plasmid pJU169 (*PRP8 URA3 CEN ARS*). Lys^+ transformants were screened for disruption of the chromosomal *PRP8* locus by their inability to grow on 5FOA-containing media. The disruption was confirmed by a whole cell PCR assay, and the 5FOA induced lethality was rescued by the presence of a non-*URA3*-marked *PRP8* plasmid (data not shown).

PCR mutagenesis: The strategy we used for mutagenic PCR is described in Figure 1 (Muhlrad et al., 1992). Mutagenic PCR conditions have been described previously (Leung et al., 1989). Primer sequences and Mn^{++} concentrations were as follows: A1 5'-GCATGCTCGAGACTTCAAAGCATGG-3', A2 5'-ACATGCGCGGATTTGATGCAT-3', .1 mM MnCl2; B1

5'-AATACAAAAGATGCGATGTCG-3', B2 5'-GCTCGCCCTAGGTTAACGTCC-3', .03 mM MnCl2; C1 5'-CAGAGATACCACCTCTTCTGG-3', C2 5'-TAGAAAATGCAGTGTACGATG-3', .05 mM MnCl2; D1 5'-TGATCGGTATCGATTTGGCT-3', D2 5'-CTAAATACATCGATTTGTTCG-3', .1 mM MnCl2. When MnCl2 was included, 200 µM dATP, 1 mM dGTP, dTTP and dCTP were used. Without MnCl2, all dNTPs were used at 1 mM.

PCR reaction volumes were performed in 4 x 50 µl aliquots using . Amplitaq polymerase (Perkin-Elmer). The samples were then extracted with phenol/chloroform and precipitated with .4 volumes 3M NaOAc (80 µl) and 1.4 volumes isopropanol (280 µl). The appropriate PCR product was cotransformed with gapped plasmid pJU225, with the PCR DNA in a mass ratio of 5 PCR DNA: 1 vector DNA. The enzymes used for gapping pJU225 were: (A) *NheI/SalI*, (B) *SalI/SpeI*, (C) *SpeI/MscI*, (D) *MscI/SphI*.

Gap Repair: To determine whether the mutations we identified were necessary and sufficient to generate the phenotypes we observed, regions of a mutant (or wild type control) plasmid were amplified using non-mutagenic PCR conditions (see above) and cotransformed with gapped wild type vector (pJU225). For *prp8-101 - prp8-107*, the amplified region spanned nucleotides 5322-6024 of the coding region and the vector was cut with *MscI* at position 5723. For *prp8-121 - prp8-125*, the amplified region spanned nucleotides 4125-5322 of the coding sequence and the vector was partially digested with *BstEII* in the presence of 100 µg/ml ethidium bromide (EtBr). The EtBr enriched for singly cut plasmids. The relevant cut site that can be repaired from the PCR DNA is at nucleotide 4829 of the coding region. The other *BstEII* site is at nucleotide 6931. In each case, the mutant PCR DNA stimulated the appearance of a mutant phenotype by 10- to 1000-fold over amplified wild type DNA.

RNA analysis: RNA preparation and primer extensions assays were performed as previously described (Lesser & Guthrie, 1993a). Results were quantitated by PhosphorImage scanning of duplicate or triplicate samples.

RESULTS

Mutagenesis strategy for *PRP8*: In order to identify new alleles of *PRP8*, we utilized PCR, which can efficiently mutagenize a selected region of DNA (Leung et al., 1989; Muhlrad et al., 1992). Since the coding region of *PRP8* is very large (7.2 kb), we divided the gene into four approximately equal-sized regions (A, B, C and D) based on convenient restriction sites (Figure 1). This division was necessary to reduce the high frequency of null alleles that would be expected if the whole coding region were mutagenized simultaneously, and it also facilitated the mapping of mutations. Each region of the gene was amplified using Taq polymerase, with and without added manganese. Manganese has been reported to increase the error rate of Taq polymerase by approximately five-fold (Leung et al., 1989). Each PCR fragment was cotransformed into a recipient strain along with an appropriately gapped, high copy *PRP8* plasmid. Since the PCR fragments contained approximately 300 nucleotides of homology on each side of the gapped region, they could direct repair of the gapped *PRP8* plasmid (Muhlrad et al., 1992; Figure 1B). When no PCR-amplified DNA is cotransformed with the gapped plasmid, the gap can be repaired efficiently from the plasmid-borne copy of *PRP8* that is already in the cell (Table 1; see below). However, based on the frequencies of null mutants obtained in the presence of PCR-amplified DNA, the PCR DNAs appear to compete strongly as repair donors, probably because they contain free ends. The PCR DNA also stimulated transformation efficiency approximately five-fold (data not shown).

The transformants were initially scored in several assays. Each recipient strain contained a low copy, *URA3*-marked plasmid bearing the wild type *PRP8* gene and a high copy, *LEU2*-marked plasmid bearing an *ACT1-CUP1* gene fusion as a splicing reporter. By selecting against cells harboring the wild type *PRP8* plasmid with 5-fluoro-orotic acid (5FOA), the recessive phenotypes of the *PRP8* mutants could be determined (Figure 1C). As a measure of mutagenic efficiency, we scored null alleles by their inability to grow on 5FOA-containing plates. Without cotransformation of PCR DNA, the null frequency was

low for each gapped region (4-8 percent). Cotransformation of PCR DNA caused a 4- to 10-fold increase in the frequency of null alleles, and the use of manganese mutagenized PCR DNA caused a further increase to a null frequency of 48%-76%. We also screened a limited number of mutants (300 total in six separate pools; see Table 1) for temperature or cold sensitivity after selection on 5FOA. Although the number of transformants scored this way was too low to be statistically significant, we found a surprisingly large number of temperature sensitive *PRP8* mutants (20/300). These were distributed among the three regions that we scored: B, C and D. We did not recover any cold-sensitive mutants. In summary, the mutagenesis strategy that we have developed efficiently targets four different subregions of *PRP8* and allows rapid identification of alleles that confer different splicing phenotypes (Table 1; see below).

Identification of new *prp8-101*-like alleles: In order to identify new alleles of *PRP8* that affect uridine tract recognition, we introduced the 8 pools of our library (A-D, +/- MnCl2) into a recipient strain harboring the reporter ACT-CUP +T PyDOWN (Figure 2A). The reporter directs synthesis of an intron-containing RNA with duplicated 3' splice sites, one uridine-rich and the other adenosine-rich. The intron is fused to the *CUP1* gene, which encodes a copper-chelating metallothionein homolog that can be used as a selectable marker (Lesser & Guthrie, 1993a). Use of the branchsite-proximal, uridine-rich 3' splice site produces a message where the initiator AUG in exon one is out of frame with the *CUP1* coding sequence, whereas use of the branchsite-distal, adenosine-rich 3' splice site results in the initiator codon being in frame with *CUP1* (Patterson & Guthrie, 1991; Umen & Guthrie, 1995a). In wild type strains, the uridine-rich 3' splice site is preferred > 20:1 over the adenosine-rich competitor (Figure 3A, lane 2; Table 2). Loss of uridine recognition and the consequent activation of the adenosine-rich 3' splice site generates more in-frame message and Cup1 fusion protein. Since the chromosomal copy of *CUP1* is deleted in this strain, we can detect increased splicing to the adenosine-rich 3' splice site by

measuring increased copper resistance (Lesser & Guthrie, 1993a; Umen & Guthrie, 1995a).

The wild type strain harboring +T PyDOWN cannot survive at copper concentrations above .05 mM while the *prp8-101* mutant allows growth at .18-.25 mM copper (Umen & Guthrie, 1995a). The pools of transformants were replica plated to copper concentrations ranging from .1 to .5 mM either before or after selection on 5FOA. No cells transformed with the A and B libraries grew at or above .1 mM in this assay, indicating that it is difficult or impossible to mutate regions A and B and obtain a *prp8-101*-like phenotype. We obtained a combined total of 45 presumptive dominant isolates that grew at or above .1 mM copper before FOA selection from cells transformed with the C and D libraries. The *prp8-101* mutation lies in an overlap region of the C and D libraries and could, in principle, be obtained from either pool (Umen & Guthrie, 1995a). After selection on 5FOA, there were approximately 100 additional presumptive recessive isolates that could grow at or above .1 mM copper (Table 1). As we were interested in the strongest alleles, we focused on twelve isolates that grew at the highest copper concentrations (.15-.25 mM) prior to 5FOA selection.

By transformation into *E. coli*, we were successful in rescuing the mutant *PRP8* plasmids from eleven of the twelve isolates. These clones, which were later found to represent seven different alleles of *PRP8*, *prp8-101* - *prp8-107*, were retransformed into the original yeast strain harboring the reporter plasmid +T PyDOWN, and each was retested for the *prp8-101*-like phenotype by growth on copper-containing media. Like *prp8-101*, each new mutant conferred increased copper resistance in the presence of the wild type *PRP8* plasmid but showed slightly higher copper resistance after selection on 5FOA (Table 2, data not shown). Thus, these alleles are all haploviable and semidominant.

In order to directly test the effect of the new alleles on alternative splice site selection, we analyzed RNA isolated from mutant and wild type strains by primer extension. In wild type cells, the ratio of branchsite-proximal uridine-rich 3' splice site

usage versus branchsite-distal, adenosine-rich 3' splice site is 26:1. In all mutant strains, splicing to the distal splice site is activated and the ratio of splice site usage is more balanced (Figure 3A, lane 2 versus lanes 3-9; Table 2). *prp8-101* and *prp8-106* display the largest change in this ratio (2.0:1 and 2.2:1). The copper resistance does not correlate strictly with the primer extension data because the mutations cause two phenotypes. The first phenotype is activation of the distal splice site, which increases copper resistance, and the second is an overall decrease in the efficiency of splicing, which tends to reduce levels of both the proximal and distal mature messages and decrease copper resistance. Thus, in *prp8-106*, the overall efficiency of splicing is slightly reduced in comparison with *prp8-101*.

We also examined splicing with a different 3' splice site competition construct, +A WT (Figure 2A; Patterson and Guthrie, 1991; Umen and Guthrie, 1995a). The RNA produced from this plasmid differs from that of +T PyDOWN in that the branchsite-proximal 3' splice site is adenosine-rich and the branchsite-distal 3' splice site is uridine-rich. The reversed order of the 3' splice sites in this construct versus +T PyDOWN allows us to rule out differences in distance or spacing preference as the cause of the phenotype. With +A WT, the change in ratios of 3' splice site usage in the mutant strains is again altered to reduce splicing to the uridine-rich 3' splice site (Figure 3B, lane 2 versus lanes 3-9; Table 2). Thus, the change in ratio of splice site usage in the mutant strains, with both +T PyDOWN and +A WT, correlates with the placement of the uridine-rich tract, and not with the relative positions of the competing splice sites. This result confirms that the new mutants are all impaired in uridine tract recognition *per se*.

Mapping uridine recognition mutations: To determine the nature and location of the mutations in these new alleles, we sequenced the region of overlap between the C and D libraries in each clone. We reasoned that this would be a likely location for the relevant mutations since it is where the original *prp8-101* mutation lies. Among the eleven isolates, there were two that contained the E1960K alteration found in *prp8-101*. There

were three additional isolates that changed this amino acid to a G. This allele is designated *prp8-102*. The other isolates all caused a change in amino acid F1834 to L or to S. Two of the alleles, *prp8-106* and *prp8-107* only cause a change in amino acid 1834, whereas *prp8-103*, *prp8-104*, and *prp8-105* cause a change at 1834 and one additional alteration (Table 7; Figure 7).

That we repeatedly identified changes in the same two amino acids suggests that the mutations affecting these positioins are sufficient to cause the *prp8-101*-like phenotype. To test this idea, we amplified the region we had sequenced from each clone using PCR under non-mutagenic conditions and cotransformed the amplified DNA with a wild type *PRP8* plasmid that had been cut with *MscI*. The *MscI* recognition site lies in the middle of the region that we amplified for each clone and stimulates gene conversion from the amplified sequences to the plasmid. Whereas amplified wild type *PRP8* DNA generated a very low frequency of *prp8-101*-like phenotypes upon cotransformation (<1%), cotransformed DNA amplified from the mutant alleles gave rise to a mutant phenotype at a high frequency (10-40%)(data not shown). Thus, the alterations in amino acids 1960 and 1834 are sufficient to generate the *prp8-101*-like phenotype. The additional changes in *prp8-103*, *prp8-104*, and *prp8-105* may slightly alter the phenotype (see Table 2) but are not necessary to confer it. Interestingly, the additional changes in these alleles occur in conserved amino acids (Figure 7) (Hodges et al., 1995). Despite this conservation, the identity of these residues must not be critical for the function of Prp8p.

Isolation of 3' splice site fidelity mutants: There is mounting evidence that Prp8p may be present at or near the active site of the spliceosome during the second catalytic step (Teigelkamp et al., 1995; Umen & Guthrie, 1995a; Umen & Guthrie, 1995b). Therefore, we wished to test whether Prp8p plays a functional role in recognition of the PyAG trinucleotide sequence at the 3' splice junction. Our strategy was to screen for alleles of *PRP8* that suppress the phenotypes of point mutations in the PyAG motif. The reporter constructs we utilized (Figure 2B, Set I) contain the sequence GAG (3' GAG) or

UUG (3' UUG) at the 3' splice junction. In addition, parts of exon two were altered to eliminate possible cryptic 3' splice sites, and in 3' UUG, two nucleotides of the intron near the 3' splice site were fortuitously deleted during its construction. 3' GAG confers a copper resistance of .1 mM; 3' UUG confers a resistance of .05 mM, versus 2 mM for a wild type 3' splice site (Table 3; Table 4).

Mutagenized *PRP8* libraries were introduced into strains harboring one of these two constructs and transformants were selected for growth on .25-1 mM copper before and after 5FOA selection against the wild type *PRP8* plasmid (Figure 1). For each construct, we isolated ~15-20 suppressors, distributed in regions B, C and D, that grew above .25 mM copper prior to 5FOA selection. The majority of suppressors arose from region C and the strongest suppressors were also from the C region. After 5FOA selection, several dozen more weaker suppressors were isolated but were not characterized further. Again, the majority of these were in the C region (Table 1).

We chose six of the strongest suppressors (three for each 3' splice site mutation), and were able to recover five of the mutant plasmids by transformation into *E. coli*. Each plasmid conferred suppression upon retransformation into yeast. These alleles are designated *prp8-121 - prp8-125*. The level of suppression on copper is very high for these mutants, reaching 15-fold over wild type for those selected against 3' UUG (*prp8-121* and *prp8-122*) and 7.5-10-fold over wild type for those selected against 3' GAG (*prp8-123-prp8-125*) (Table 3). Each of the alleles is dominant and haploviable (data not shown).

Specificity for suppression of 3' splice site mutations: We first tested whether these alleles could suppress mutations in other parts of the intron and whether they altered 3' splice site uridine tract recognition, similar to *prp8-101-prp8-107*. We examined splicing of constructs containing a mutation in the branch nucleotide from A to C (A259C) and a construct containing a G to A mutation at the fifth position of the intron (G5A). For G5A, we measured splicing to both the authentic 5' splice site and to an upstream cryptic 5' splice site at position -5, which is activated by the G5A mutation (Parker & Guthrie, 1985;

Lesser & Guthrie, 1993b). Finally, we measured splicing in the 3' splice site competition construct +T PyDOWN. We assayed splicing of these *ACT1-CUP1* fusions by growth on copper-containing media.

Splicing of A259C is unaffected in *prp8-124* and decreased two-fold in *prp8-121, -122, -123,* and *-125* (Table 4). Normal splicing of G5A is unaffected in strain *prp8-123* and decreased between two and greater than five-fold in *prp8-121*, *prp8-122* and *prp8-124*. We see a very slight suppression in *prp8-125* (less than two-fold). For aberrant splicing of G5A, we see a similar pattern except that *prp8-125* splices to the aberrant site at a similar level as wild type and *prp8-123* slightly suppresses aberrant splicing. Again, this suppression (less than two-fold) is very slight compared with the suppression this allele confers to introns with PyAG mutations (Table 4). Finally, uridine tract recognition is unaffected by *prp8-121-prp8-125* since +T PyDOWN splicing is unchanged in these strains (Table 4). In summary, the strong suppression phenotype of alleles *prp8-121 - prp8-125* appears to be specific for PyAG mutations at the 3' splice site. Other alterations in the intron are either relatively unaffected or exacerbated by these alleles.

Allele specificity for different PyAG mutations:

Construct set I (3' GAG and UUG): As a first test of allele specificity for different PyAG alterations, we transformed *prp8-121* and *prp8-122* into a strain harboring the 3' GAG construct, and *prp8-123-prp8-125* into a strain harboring the 3' UUG construct. This experiment revealed that either set of alleles could suppress a different 3' splice site mutation other than the one it was selected against, but less efficiently. This observation was confirmed by directly analyzing splicing in mutant and wild type strains by primer extension (Figure 4A and 4B, lane 2 versus lanes 3,4 and 5-7; Table 3). In this and in subsequent primer extension experiments, we determined the ratio of mRNA (M) to lariat intermediate (LI) as a measure of suppression. This ratio was used because 3' splice site mutations in yeast affect only the second catalytic step, and the M/LI ratio represents the overall rate of the second catalytic step *in vivo* (Pikielny & Rosbash, 1985; Frank &

Guthrie, 1992). By copper and primer extension assays, *prp8-121 -prp8-125* suppress both types of 3' splice site mutations but exhibit some allele preference (Table 3).

Construct set II (3' UAG, AAG, GAG and UGG): To better characterize the effects of *prp8-121-prp8-125* on 3' splice site utilization, we tested a more extensive set of PyAG point mutations. These include the changes UAG to AAG, GAG, UGG, UUG, UAU and UAC introduced in the context of the standard *ACT1-CUP1* gene fusion. The exon sequences of Construct Set II are different from those of Set I in the region between the 3' splice site and the *CUP1* coding sequence (Figure 2C, Set II versus Set I). These were transformed into wild type strains and *prp8-121 - prp8-125* mutatants and splicing was analyzed by primer extension and/or growth on copper-containing media (Tables 5 and 6). We divided the reporter constructs into two groups based on the severity of splicing defect that each causes, and on the pattern of splicing each exhibits. For 3' UAG (Wild type), AAG, GAG and UGG, we observe splicing to only the correct 3' splice site in both wild type and mutant strains, and observe growth on copper-containing media above background levels. For 3' UAG and AAG the levels of lariat intermediate produced are too low to accurately determine the (M/LI) ratio. For these two constructs, we only measured growth on copper-containing media. For 3' GAG and UGG, we measured splicing by growth on copper-containing media and by primer extension.

For the wild type *ACT1-CUP1* intron (3' UAG), we observe a slight reduction in splicing efficiency in *prp8-121 - prp8-125* strains compared with the wild type strain (Figure 5A, lane 2 versus lanes 3-7, Table 5). 3' AAG is spliced equally well in wild type strains and in *prp8-121 - prp8-125* (Figure 5B, lanes 2-7; Table 5). The splicing block caused by the 3' GAG construct is substantially suppressed in *prp8-121 - prp8-125* strains. By growth on copper-containing media, the pattern of allele specificity is similar to that seen for 3' GAG in construct set I, but the differences between *prp8-123-prp8-125* and *prp8-121 -prp8-122* are less pronounced (Table 5). This difference is also reflected in the primer extension analysis (Figure 5C, lane 2 versus lanes 3,4 and 5-7; Table 5). The

exceptions are *prp8-121* and *prp8-123*, which have primer extension M/LI values that are higher and lower, respectively, than expected from growth on copper assays (Table 5). Since growth on copper-containing media itself is a reproducible measure of the amount of mature message produced, these results suggest that the mutants may also alter the stability of the lariat intermediate as a secondary phenotype and thus slightly skew the M/LI ratio. Although we include an internal standard (U1 snRNA labeled "control") in each primer extension experiment, the reproducibility between samples for U1 or other standards is approximately 2- to 3-fold, whereas the M/LI ratio is highly reproducible between samples (usually less than 1.5-fold). Thus, the error introduced in trying to measure the amount of lariat intermediate (LI) for each sample would be greater than the differences we wished to measure.

For 3' UGG we see a wide range of suppression on copper varying from two-fold in *prp8-122* to ten-fold in *prp8-123* (Table 5). The results of primer extension analysis generally match the results from growth on copper assays but, again, vary somewhat due to the relative contributions of decreased lariat intermediate and increased mature message to the (M/LI) ratio (Figure 5D; Table 5). In summary, for wild type and 3' AAG, which display little or no splicing defect, we see no suppression. For 3' splice site alterations that exhibit a strong splicing defect (3' GAG and UGG), we see suppression by *prp8-123* - *prp8-125*.

Construct set II (3' UUG, UAU and UAC): With 3' U<u>U</u>G, UA<u>U</u> and UA<u>C</u>, we observe splicing to the correct 3' splice and to an upstream, cryptic 3' splice site at position -5 in the intron. The sequence preceding the cryptic cleavage site is AUG/, as has been noted previously (Parker & Siliciano, 1993). These constructs produce very little mature message in a wild type strain and show growth on copper-containing media at or near background levels. Although we observe increased resistance to copper in *prp8-121* - *prp8-125* with 3' U<u>U</u>G, UA<u>U</u> and UA<u>C</u> (data not shown), the degree of suppression is difficult to assess by growth on copper-containing media since the starting resistances are

not measurable and because the cryptic 3' splice site is not in-frame with the *CUP1* coding sequence. Therefore, we used only primer extension to quantitate splicing of these constructs. We measured three values: the efficiency of splicing to the wild type 3' splice site (M_{wt}/LI), the efficiency of cryptic 3' splice site usage (M_{crypt}/LI) and the ratio of authentic to cryptic splice site usage ($M_{wt}/M_{cryptic}$).

The pattern of suppression of 3' UUG in construct set II (M_{wt}/LI) is generally similar to what we observed for 3' UUG of construct set I, with differences between alleles being less pronounced. The unexpectedly weak apparent suppression by *prp8-122* (3-fold increase in M/LI) in this context may reflect an increase in the stability of the lariat intermediate with this construct. In general, the allele specificity has been preserved for the two sets of constructs containing 3' GAG and 3' UUG alterations, but the levels of suppression are also clearly subject to modification by surrounding sequence context.

The overall pattern we see for 3' UUG, 3' UAU and 3' UAC in *prp8-121* - *prp8-125* is suppression of both the authentic and cryptic splice sites in the mutant strains (Figure 6A-6C, lane 2 versus lanes 3-7; Table 6). The authentic 3' splice site is usually suppressed slightly better than the cryptic 3' splice site but not always. For example, the *prp8-122* strain suppresses cryptic splicing of 3' UUG and 3' UAU better than it suppresses authentic splicing (Figure 6A and 6B, lane 2 versus 4; Table 6). With 3' UAC, suppression values in Table 6 appear lower than what is visualized in the primer extension experiment (Figure 6C, lane 2 versus lanes 3-7) because the wild type sample (lane 2) is underloaded and because the mutants may stabilize the lariat intermediate (LI) compared to wild type.

The ratio of authentic to cryptic splice site usage ($M_{wt}/M_{cryptic}$) is useful for evaluating the *prp8-121* - *prp8-125* strains since it gives a measure of relative preference for two competing 3' splice sites. This ratio is altered substantially in some of the mutant strains. For example the ratio changes four-fold in favor of the authentic 3' splice site with 3' UAC in a *prp8-123* strain. In contrast, ($M_{wt}/M_{cryptic}$) changes two-fold in favor of the

cryptic 3' splice site in *prp8-122* with 3' UUG (Table 6). In general, the deviations in ($M_{wt}/M_{cryptic}$) for the mutant strains support the notion that these alleles cause altered recognition of the 3' splice site.

3' splice site alteration in a heterologous intron: Because surrounding sequence context has some effect on suppression (cf 3' UUG and 3' GAG; Tables 3, 5 and 6), we wished to determine whether the 3' splice site suppression we observed with actin intron constructs would also be observed in the novel context of a different intron. We utilized an *rp51a-lacZ* fusion construct with a mutant 3' splice site (UGG) to address this question (Chanfreau et al., 1994). The splicing of this construct is suppressed between 1.2- and 15-fold when assayed in *prp8-121 - prp8-125* strains, indicating that suppression is not intron specific (Table 4). Thus, these alleles of *PRP8* do not require a specific surrounding sequence context to suppress PyAG alterations at the 3' splice site.

In summary, *prp8-121 - prp8-125* demonstrate several interesting properties: (i) They suppress the effects of point mutations in the 3' splice site but not the effects of mutations elsewhere in the intron. (ii) 3' splice site suppression is not specific for any particular point mutation in the PyAG motif, but the alleles display a complex pattern of preferences. (iii) The preferences displayed by the alleles preclude ordering them into an allelic series based on relative strength of suppression. (iv) The suppression phenotype can be partially altered by surrounding sequence context, as evidenced by differences between construct set I and II. (v) However, a specific sequence context is not necessary for suppression.

Mapping 3' splice site fidelity mutations: In order to determine the identity of the mutations in *prp8-121 - prp8-125* , we first used gap repair to map the region of *PRP8* that contained the relevant change. We then sequenced a 1.2 kb *BstBI-ClaI* fragment from region C that was predicted to contain the relevant alteration for each allele. In each case, this region was found to contain one or more missense mutations (Table 7).

prp8-122 and *prp8-124* have single alterations in this fragment whereas *prp8-121*, *-123* and *-125* encode multiple changes.

To determine whether or not these changes were sufficient to elicit the observed suppression phenotype, we used a similar strategy as that employed for the *prp8-101*-like alleles. The cloned *1.2 kb BstBI-ClaI* fragment from each allele and from wild type was amplified using PCR under non-mutagenic conditions. This DNA was then cotransformed with a *PRP8* plasmid that had been partially cut with *BstEII*. One of two *BstEII* sites in the *PRP8* gene lies in the middle of the cloned region (see Methods). In each case, the cotransformed PCR fragment from the mutant allele generated a suppressor phenotype at a rate >100-fold higher than that of cotransformed wild type PCR DNA. The *prp8-124* PCR DNA generated suppressors at a frequency only ten-fold above wild type, but this result was expected since the mutation on this fragment lies much farther from the *BstEII* site than the changes in the other alleles.

Since *prp8-121*, *-123* and *-125* each contain more than one mutation, it was of interest to know which alteration(s) is responsible for the suppression phenotype. We took advantage of the fact that there is at least one mutation in each of these alleles on either side of the aforementioned *BstEII* site. Reciprocal swaps were constructed that contained all six possible wild type/mutant or mutant/wild type combinations of parent vector and heterologous 2.1 kb *BstEII* fragment. In each case, a single mutation was isolated and found to be responsible for the phenotype (Table 7; Figure 7). Four of the suppressor mutations alter amino acids in a region of 44 residues spanning positions 1565-1609. The fifth, in *prp8-124* , alters an amino acid 166 residues N-terminal to this stretch.

DISCUSSION

Several studies have shown that Prp8p interacts with intron sequences that are critical for splicing (Wyatt et al., 1992; Teigelkamp et al., 1995; Umen & Guthrie, 1995a).

Furthermore, the large size (280 kd) and sequence conservation of Prp8p suggest that it mediates multiple interactions with spliceosomal RNAs and proteins (Anderson et al., 1989; Pinto & Steitz, 1989; Hodges et al., 1995). However, it has been difficult to assess the functional significance of these interactions. Removal of Prp8p by heat inactivation of a temperature sensitive allele or by *in vivo* depletion reveals a role for the protein in maintaining the integrity of the U4/U5/U6 triple snRNP (Brown & Beggs, 1992). Since the U4/U5/U6 snRNP is required during spliceosome assembly, however, these methods cannot be used to assess the function(s) of Prp8p within the spliceosome.

Uridine tract recognition by Prp8p: By repeating our original search for uridine tract recognition mutants using a mutagenized *PRP8* library, we have isolated six new alleles of *PRP8* that contain alterations in one of two amino acids, E1960 and F1834. As mutations in these amino acids were isolated multiple times, it appears that we have saturated this search, at least for single changes that yield a strong uridine tract recognition phenotype. Notably, none of the new alleles isolated are significantly stronger in phenotype than *prp8-101*, and all are haploviable. It is possible that alleles of *PRP8* exist which change the ratio of splice site usage even more than the ones we isolated, but do not result in a net increase in usage of the adenosine-rich 3' splice site. However, one observation argues that the alleles we isolated are the most severe that could be recovered from this screen. The ratios of proximal to distal splice site usage for +T PyDOWN in our mutant strains is similar to the ratio of splice site usage in a wild type strain when no uridine tract precedes either 3' splice site in a cis competition (Patterson & Guthrie, 1991). Thus, *prp8-101-prp8-107* behave as though there were no uridine tract preceding the proximal 3' splice site in +T PyDOWN.

It is clear from this work that only a very limited and specific set of changes in *PRP8* leads to loss of uridine preference. Residue 1834 contains a conservative change from F to Y in *C. elegans*, and residue 1960 is conserved as an E in *S. cerevisiae* , *C. elegans*, *H. sapiens*, and *O. sativa* (Hodges et al., 1995) (Figure 7). It is important to bear

in mind that we selected and characterized only the strongest uridine recognition alleles of *PRP8*. The weaker alleles that we did not characterize were also isolated from the C and D libraries and probably lie in the same region of overlap. Supporting this idea, changes in positions 1880, 1922 and 1946 occur along with changes in position 1834 in *prp8-103-prp8-105*, and these additional changes cause some alteration in the phenotype (Tables 2 and 7). Thus, the region encompassing residues 1834-1960 of Prp8p is likely to be involved in uridine tract recognition. Until a larger number of weaker alleles are mapped, we cannot precisely define the limits of this functional domain. Nonetheless, we have identified the key amino acids in Prp8p, E1960 and F1834, that make the largest contribution to uridine tract recognition. These residues are predicted to function either through direct contact with the uridine tract or by stabilizing a structure in Prp8p that binds this sequence. Supporting the idea that residue 1960 facilitates uridine tract binding by Prp8p is our observation that the E1960K mutation, *prp8-101*, causes a decrease in crosslinking of Prp8p to the 3' splice site *in vitro* (Umen & Guthrie, 1995b).

Prp8p governs the fidelity of 3' splice site usage: Because Prp8p might be present at or near the active site during the second catalytic step, we wished to determine whether this protein is involved in recognition of the PyAG nucleotides that directly precede the reactive phosphate at the 3' splice junction. The PyAG motif is known to be critical for proper 3' splice site recognition and utilization (Reed & Maniatis, 1985; Ruskin & Green, 1985; Vijayraghavan et al., 1986; Fouser & Friesen, 1987; Reed, 1989; Parker & Siliciano, 1993; Chanfreau et al., 1994). In this work we have identified five new alleles of *PRP8* that suppress the inhibitory effects of PyAG alterations at the 3' splice site. This phenotype markedly contrasts that displayed by the uridine tract recognition mutant *prp8-101*, which strongly exacerbates the splicing defect of PyAG alterations (Umen & Guthrie, 1995a). Moreover, the PyAG suppressor alleles *prp8-121* - *prp8-125* do not show a defect in uridine tract recognition, nor do they suppress alterations at the branchsite.

Three of the five alleles, *prp8-121*, *-122* and *-124* actually exacerbate the effects of mutations at both the branchsite and 5' splice site. *prp8-123* and *prp8-125* have no effect on splicing of the branchsite mutant we tested, and with a 5' splice site mutation, show very weak suppression of either aberrant or normal 5' splice site cleavage. It is possible that this weak 5' splice site suppression results from the secondary mutations associated with *prp8-123* and *prp8-125*. In any case, weak 5' splice site suppression is not obligatorily coupled to 3' splice site suppression since other 3' splice site suppressor mutants do not display this phenotype. Finally, the suppressors function in the context of a heterologous intron with a 3' splice site mutation. Thus, these alleles are specific in their suppression of PyAG mutations and functionally distinct from the alleles which affect uridine tract recognition (Table 4).

prp8-121 - *prp8-125* were selected using two different PyAG variants, GAG and UUG (Table 3). In principle, suppression of 3' splice site mutations by these alleles of *PRP8* could occur by an indirect mechanism. For example, if these alleles alter a kinetic parameter that affects fidelity of 3' splice site usage, we might expect to see similar suppression for all 3' splice site mutants, or an allelic series where some alleles always generate higher levels of suppression than others. An example of an indirect mechanism is suppression of branchsite alterations by mutations in *PRP16*, which encodes an RNA-dependent ATPase (Burgess et al., 1990; Schwer & Guthrie, 1991). The *PRP16* suppressor alleles affect the ATPase activity of the protein, and there is a correlation between strength of suppression and reduction in ATPase activity (Burgess & Guthrie, 1993). The mechanism of suppression is indirect, in the sense that it does not require altered recognition of aberrant branchsites by Prp16p. Instead, the slowed ATPase activity of the mutant proteins allows more time for aberrant branch lariats, which would otherwise be degraded, to continue productively in the splicing pathway (Burgess & Guthrie, 1993).

In contrast to branchsite suppressor alleles of *PRP16*, the 3' splice site suppressor alleles of *PRP8* cannot be placed into an allelic series based on strength of suppression.

Instead, these alleles show a complex pattern of specificity with different 3' splice sites. Most illustrative of this specificity are the two 3' splice site alterations with which these *PRP8* alleles were selected, 3' GAG and 3' UUG. *prp8-121* and *prp8-122* suppress the 3' UUG alteration much better than they suppress 3' GAG, and *prp8-123*, *-124* and *-125* suppress 3' GAG much better than they suppress 3' UUG (Table 3). For other constructs, the matrix of suppression is more complicated and does not form a predictable pattern. Thus, the suppression we observe is difficult to explain via a completely indirect mechanism.

If these *PRP8* alleles do not act indirectly, how might they alter the fidelity of 3' splice site selection? The simplest possibility is that Prp8p makes direct contact with the PyAG trinucleotide and surrounding sequences to form part of a 3' splice site binding site. This idea is supported by the fact that Prp8p can be crosslinked to the 3' splice site during the second catalytic step (Teigelkamp et al., 1995; Umen & Guthrie, 1995b). Mutant Prp8p would be predicted to bind altered 3' splice sites with less discrimination than the wild type protein and allow these substrates to be utilized at a higher rate.

It is especially intriguing that mutations in two spliceosomal snRNAs, U2 A25 and U6 G52 also suppress PyAG alterations (Lesser & Guthrie, 1993b; Madhani & Guthrie, 1994). These two nucleotides are thought to participate in a tertiary interaction that forms part of the active site for the second catalytic step (Madhani & Guthrie, 1994). Moreover, the 3' splice site suppression we see for *prp8-121* - *prp8-125* appears to be stronger and more specific than that seen for the snRNA mutations. Therefore, if *prp8-121* - *prp8-125* do not alter a direct interaction between Prp8p and the 3' splice site, then these alleles are likely to alter the geometry of the active site in such a way as to relax the specificity normally imposed on 3' splice site nucleotide identity. This kind of phenotype has been previously documented for active site mutations in alpha-lytic protease which broaden substrate specificity. These alpha-lytic protease mutations cause increased flexibility in the

active site which allows a larger spectrum of amino acid side chains to be accommodated (Bone et al., 1991).

Notably, the alterations in *PRP8* that lead to strong 3' splice site suppression all cluster in a relatively small region of the protein. The changes we identified are also in residues that are conserved between yeast and worms. Four of the five alterations are in a 44 amino acid stretch, and the fifth change lies 166 residues towards the N-terminus from this stretch (Figure 7). The clustering suggests that this region of the protein may make critical contacts with the 3' splice site or active site residues. Since we found weaker suppressors in other regions of the protein, other types of structural perturbations may also alter the interaction of Prp8p with the 3' splice site. However, it should be emphasized that 3' splice site suppressors appear at a relatively low frequency. Moreover, general loss of function mutations, such as the temperature sensitive mutants we isolated in this study, do not affect 3' splice site utilization (Table 1; Umen and Guthrie, 1995a).

The two types of 3' splice site utilization mutants we isolated in this study are located relatively close to one another within the primary sequence of Prp8p (Figure 7). It is, therefore, tempting to speculate that this region of the protein functions in 3' splice site selection. However, structural data and more extensive genetic mapping are required to determine whether this 561 amino acid region of the protein also carries out other functions. In any event, our findings should provide a useful guide for future analysis aimed at a more detailed biochemical and structural understanding of how Prp8p interacts with the 3' splice site.

The mutagenized *PRP8* library we have constructed, in combination with sensitive genetic assays, has proven useful in dissecting the roles of Prp8p in the spliceosome and also in defining specific functional domains. Although we have focused our studies on 3' splice site selection and the second catalytic step of splicing, this library has many other potential uses. For example, it is now being employed to find alterations that influence the fidelity of 5' splice selection (Amy Kistler, personal communication). The library could

also be used to examine the interactions of Prp8p with snRNAs and spliceosomal proteins. Recently, we have put *PRP8* under the control of a regulated promoter to enable a screen for dominant negative alleles. This class of allele will be particularly useful in identifying new functions for Prp8p.

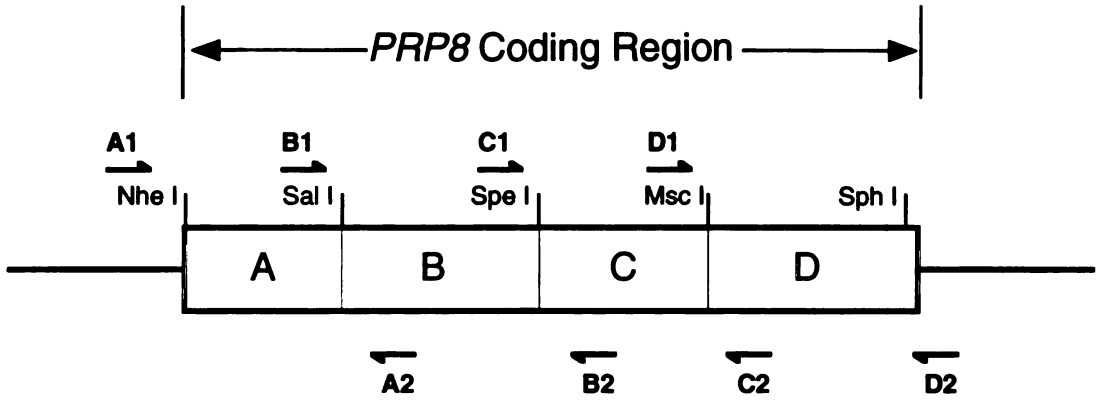
ACKNOWLEDGMENTS

We thank Cathy Collins, Ira Herskowitz, Amy Kistler and Pratima Raghunathan for critical comments on this manuscript. We are grateful for the excellent technical assistance provided by Lucita Esperas, Carol Pudlow and Heli Roiha. We thank Amy Kistler for the G5A reporter constructs used for screening mutants.

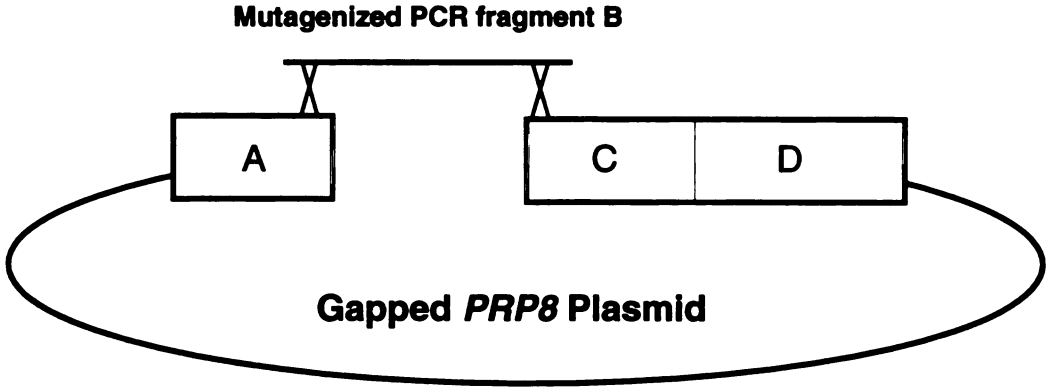
This work was supported by a NIH training grant and a NSF pre-doctoral fellowship to J.G.U., and NIH research grant GM21119 to C.G. C.G. is an American Cancer Society Research Professor of Molecular Genetics.

Figure 1. Strategy for isolation of novel *PRP8* mutants. (A) Schematic of *PRP8* gene with locations of PCR primers (depicted as arrows) and restriction sites in their approximate locations (see Materials and Methods for details). (B) Example of *in vivo* gap repair used to generate mutant library from PCR fragment B. The ends of this fragment are shown recombining with a gapped *PRP8* plasmid cut with *SalI* and *SpeI*. (C) Depicted is a yeast cell containing a reporter construct, wild type *PRP8* gene, and selectively mutagenized *PRP8* gene (region B from above). The strain is deleted for its chromosomal copy of *PRP8* and *CUP1*. *PRP8* mutants that affect splicing of the reporter construct can be selected before or after loss of the wild type *PRP8* allele by 5FOA selection.

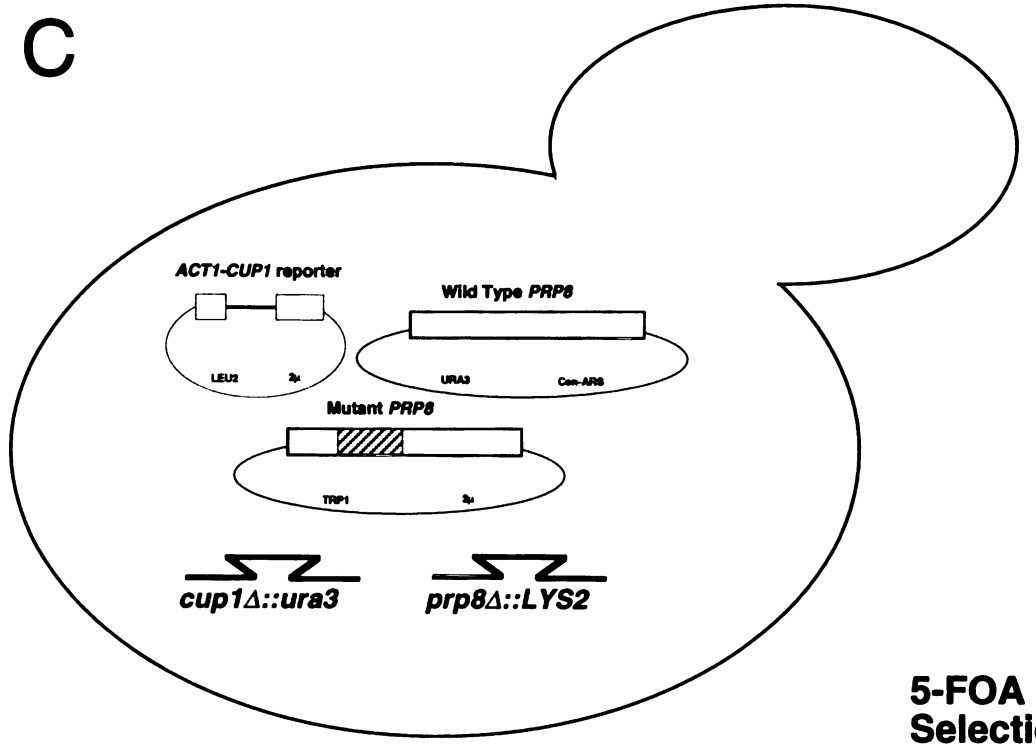
A


PRP8 Coding Region
A1
B1
C1
D1
Nhe I
Sal I
Spe I
Msc I
Sph I
A
B
C
D
A2
B2
C2
D2

B


Mutagenized PCR fragment B
A
C
D
Gapped PRP8 Plasmid

C



5-FOA Selection

Copper Selection

Dominant Mutants

Copper Selection

Recessive Mutants

Figure 2. 3' splice site reporter constructs. (A) 3' splice site competition constructs. The sequence between the branchsite (raised A in bold) and two competing 3' splice sites (underlined and bold AGs) is shown. (B) Reporter set I. As above with mutated splice acceptor sequence in bold and cleavage site followed by a / symbol. Exon 2 sequences which have been altered are also shown. (C) Reporter set II. As in (B) except only the wild type version is shown in its entirety with mutants at the 3' acceptor sequence drawn below. The cryptic UG splice acceptor is also underlined.

A

3′ Splice Site Competitions

UACUA**A**CAUCGUUUGUUGUUUCGAUU**AG**CUUCAUUCUUUUUGUUGCUAUAAAAAAAAUUU**AG** +T PyDOWN

UACUA**A**CAUCGAAACAACAAACGAUU**AG**CUUCAUUCUUUUUGUUGCUAUAUUAUAUGUUU**AG** +A WT

B

3′ Splice Site Mutants Set I

UACUA**A**CAUCGAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUAUGUU**GAG**/AAUAUUCAUUCUCCGAA 3′ GAG

UACUA**A**CAUCGAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUAUG**UUG**/AAUAUUCAUUCUCCGAA 3′ UUG

C

3′ Splice Site Mutants Set II

UACUA**A**CAUCGAUUGCUUCAUUCUUUUUGUUGCUAUAUUAUA**UG**UU**UAG**/AUGUUGGUACCGGGUUCAGCGAA

AAG

GAG

UGG

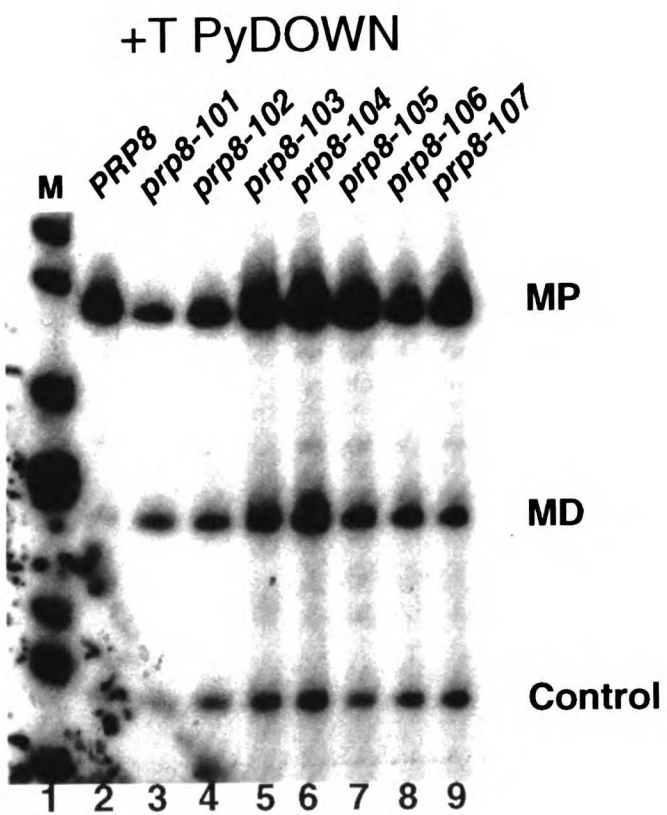
UUG

UAU

UAC

Figure 3. Splicing of 3' splice site competition constructs. (A) Primer extension analysis of +T PyDOWN splicing in wild type (Lane 2) and *prp8-101* - *prp8-107* strains (Lanes 3-9). Primer extension products corresponding to mature message from use of the branchsite proximal (MP) and branchsite distal (MD) 3' splice sites are marked next to the bands. A U1 snRNA primer extension product (control) is used as an internal control. Lane 1 (marked M) contains size markers (HpaII digested pBR325). (B) Splicing of +A WT. Lanes and primer extension products are marked as in (A). The internal control band is not shown.

A


+T PyDOWN
M
PRP8
prp8-101
prp8-102
prp8-103
prp8-104
prp8-105
prp8-106
prp8-107
MP
MD
Control
1 2 3 4 5 6 7 8 9

B

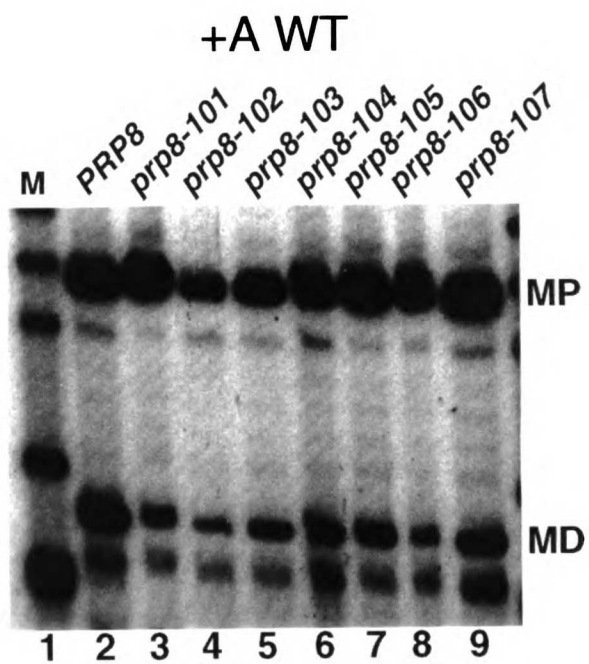

+A WT
M
PRP8
prp8-101
prp8-102
prp8-103
prp8-104
prp8-105
prp8-106
prp8-107
MP
MD
1 2 3 4 5 6 7 8 9

Figure 4. Splicing of 3' splice site mutant construct set I. (A) Primer extension analysis of splicing to 3' splice site mutant UUG in wild type (lane 2) and *prp8-121 - prp8-125* (lanes 3-7) strains. Primer extension products corresponding to mature message (M) lariat intermediate (LI) and U1 snRNA internal control (control) are labeled next to gel. Other bands are strong stops which occur variably in different lanes and experiments. Lane 1 (marked M) contains size markers (*HpaII* -digested pBR325). (B) Primer extension analysis of splicing to 3' splice site mutant GAG. Lanes and primer extension products are marked as in (A).

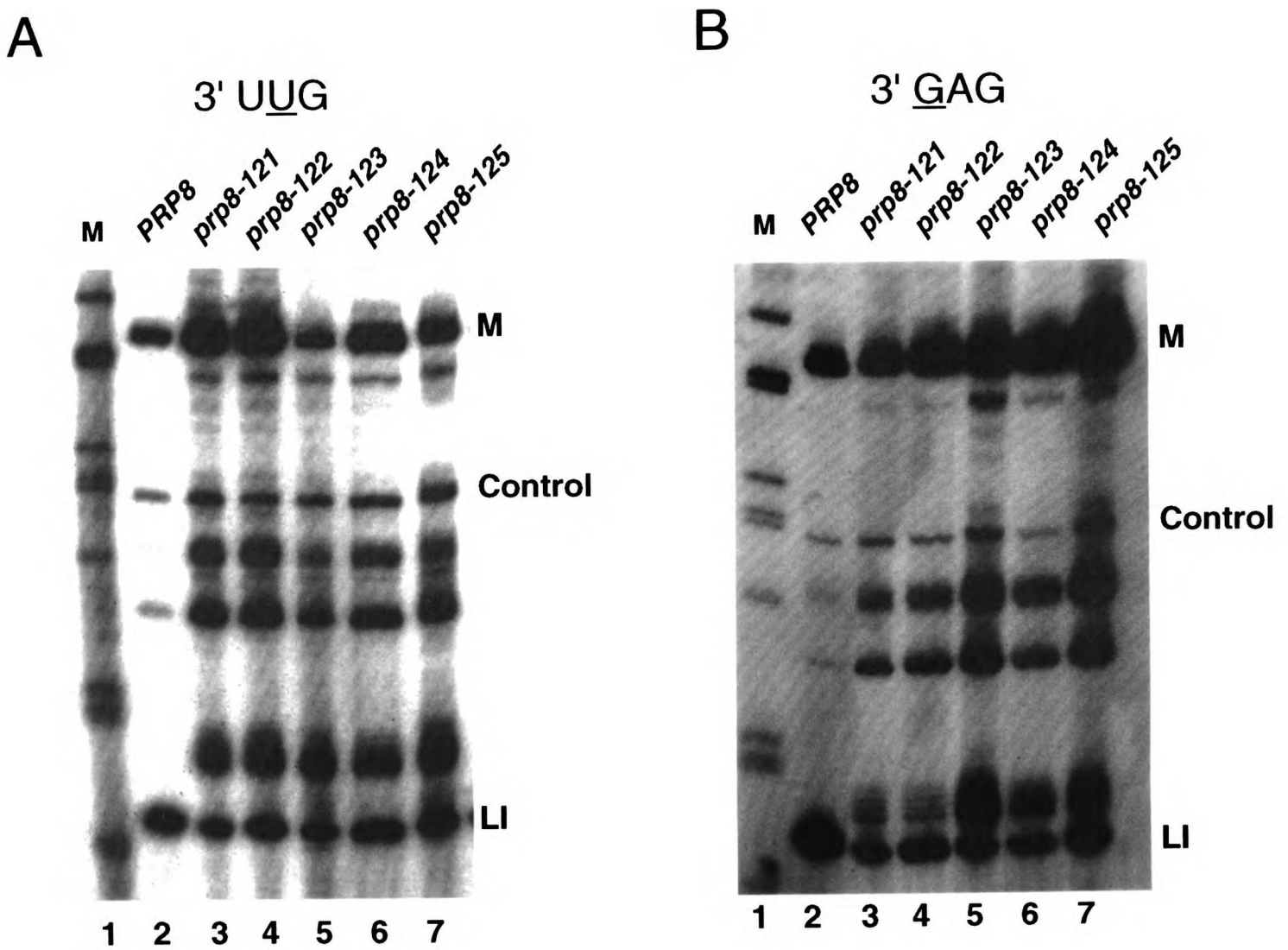
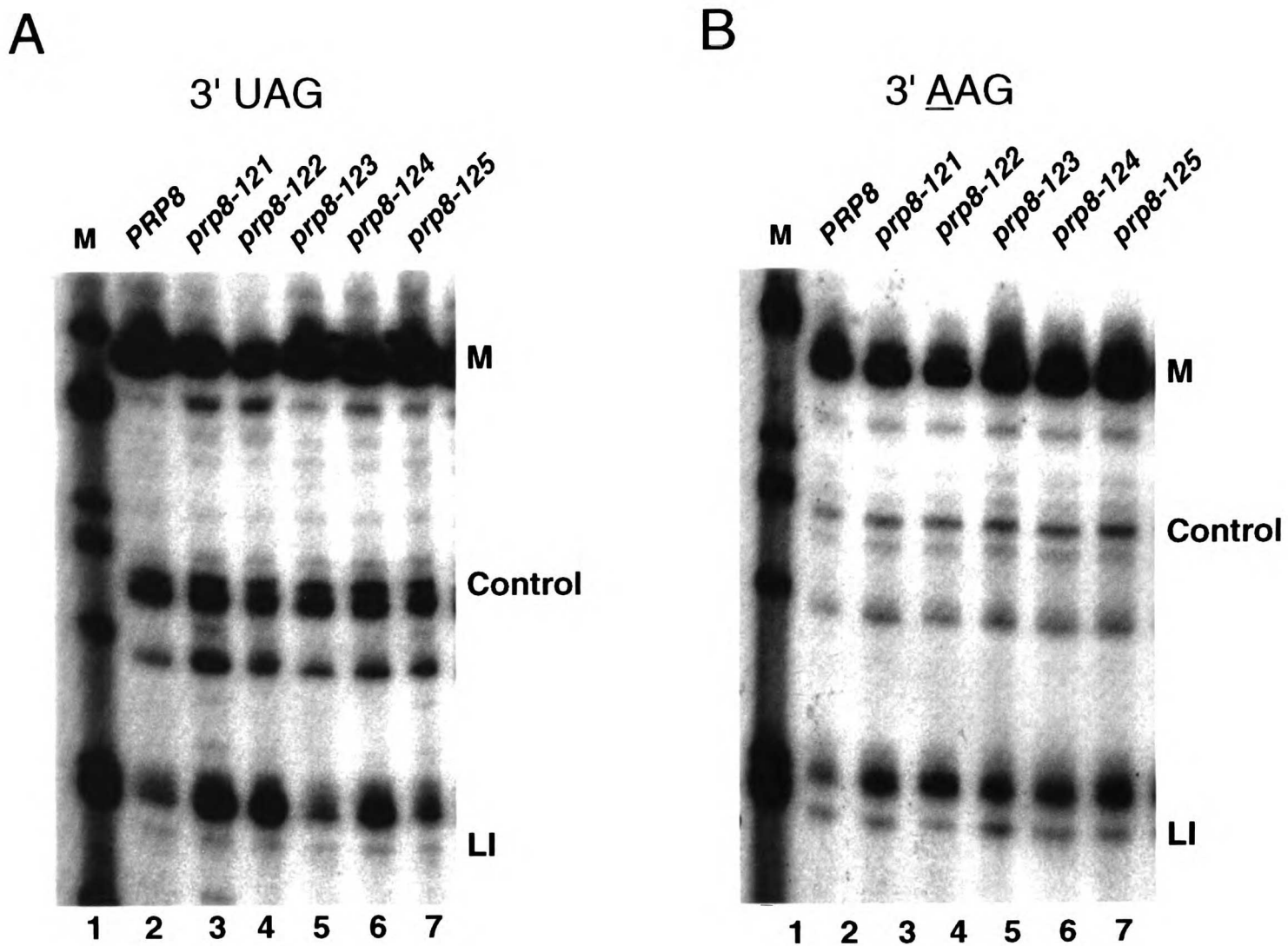

A
3' UUG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M
Control
LI
1 2 3 4 5 6 7
B
3' GAG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M
Control
LI
1 2 3 4 5 6 7

Figure 5. Splicing of 3' splice site mutant construct set II. (A-D) Primer extension **analysis** of splicing to 3' UAG (wt intron), AAG, GAG and UGG. Gels are marked as in **Figure** 4.


A
3' UAG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M
Control
LI
1 2 3 4 5 6 7
B
3' AAG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M
Control
LI
1 2 3 4 5 6 7

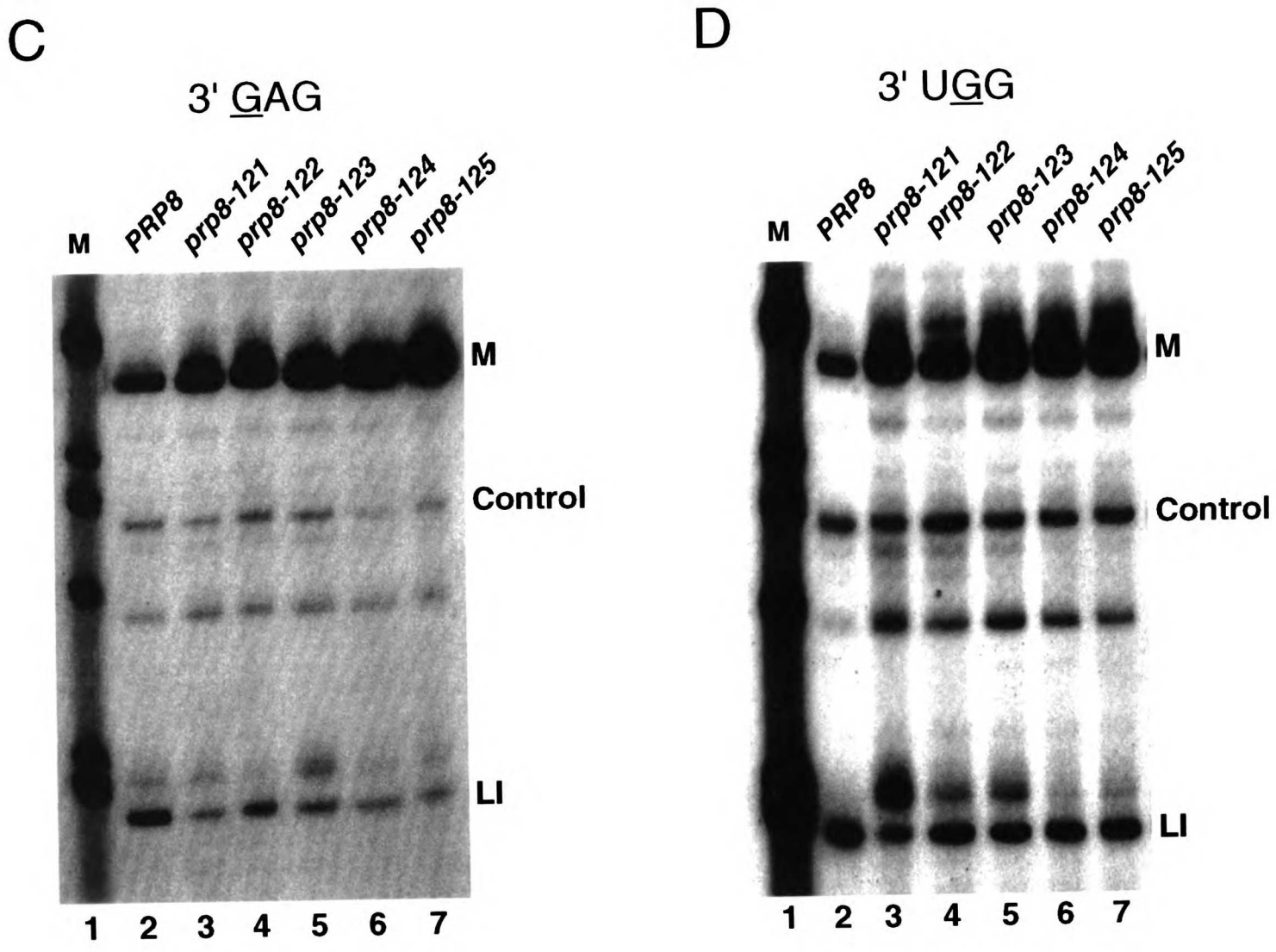

C
3' GAG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M
Control
LI
1 2 3 4 5 6 7
D
3' UGG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M
Control
LI
1 2 3 4 5 6 7

Figure 6. Splicing of 3' splice site mutant construct set II. (A) Primer extension analysis of splicing to 3' splice site mutant UUG in wild type (lane 2) and *prp8-121 - prp8-125* (lanes 3-7) strains. Primer extension products corresponding to mature message from use of the authentic 3' splice site (M_{WT}) or cryptic 3' splice at position -5 ($M_{Cryptic}$), lariat intermediate (LI) or U1 snRNA internal control (control) are marked next to the gel. Lane 1 (marked M) contains size markers (*HpaII*-digested pBR325). (B-C) Primer extension analysis of splicing to 3' UAU and UAC. Lanes and primer extension products are marked as in (A).

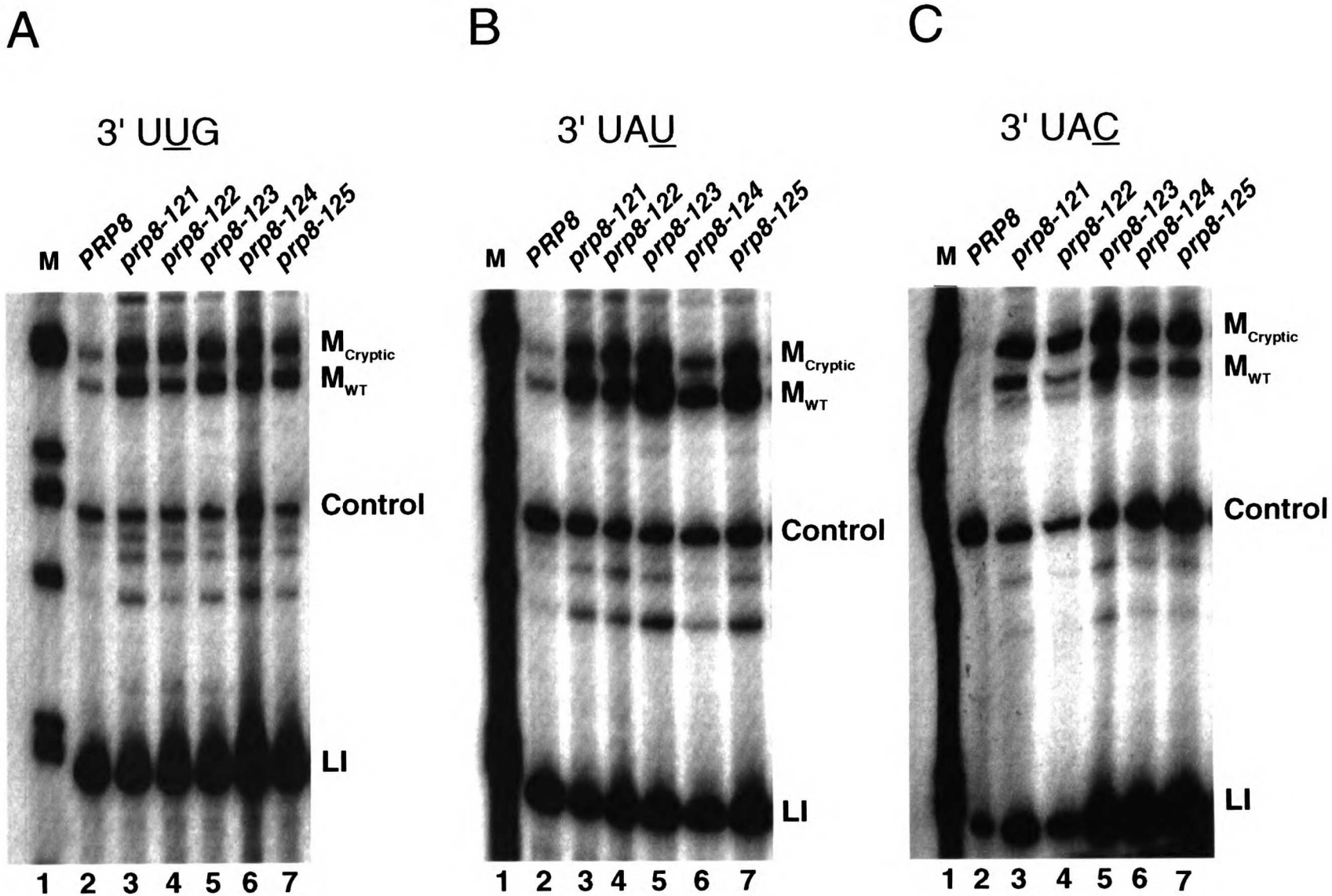

A
3' UUG
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M Cryptic
M WT
Control
LI
1 2 3 4 5 6 7
B
3' UAU
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M Cryptic
M WT
Control
LI
1 2 3 4 5 6 7
C
3' UAC
M
PRP8
prp8-121
prp8-122
prp8-123
prp8-124
prp8-125
M Cryptic
M WT
Control
LI
1 2 3 4 5 6 7

Figure 7. Domains of *PRP8* involved in uridine tract recognition and in maintaining the fidelity of 3' splice site usage. The bottom shows a schematic of Prp8, length 2413 amino acids. The hatched box shows the region where 3' splice site fidelity mutants are located and the gray box shows the location of uridine tract recognition mutants. These regions are expanded to show the location of the relevant change for each allele (see Table 7 for locations of other alterations). The nature of the mutation for each allele is shown in bold above the wild type residue. Included is an alignment with portions of *PRP8* homologs from other species: *C. elegans* (worm), *H. sapiens* (human) and *O. Sativa* (rice) (Hodges et al., 1995). The numbering on the right reflects the position in the yeast sequence.

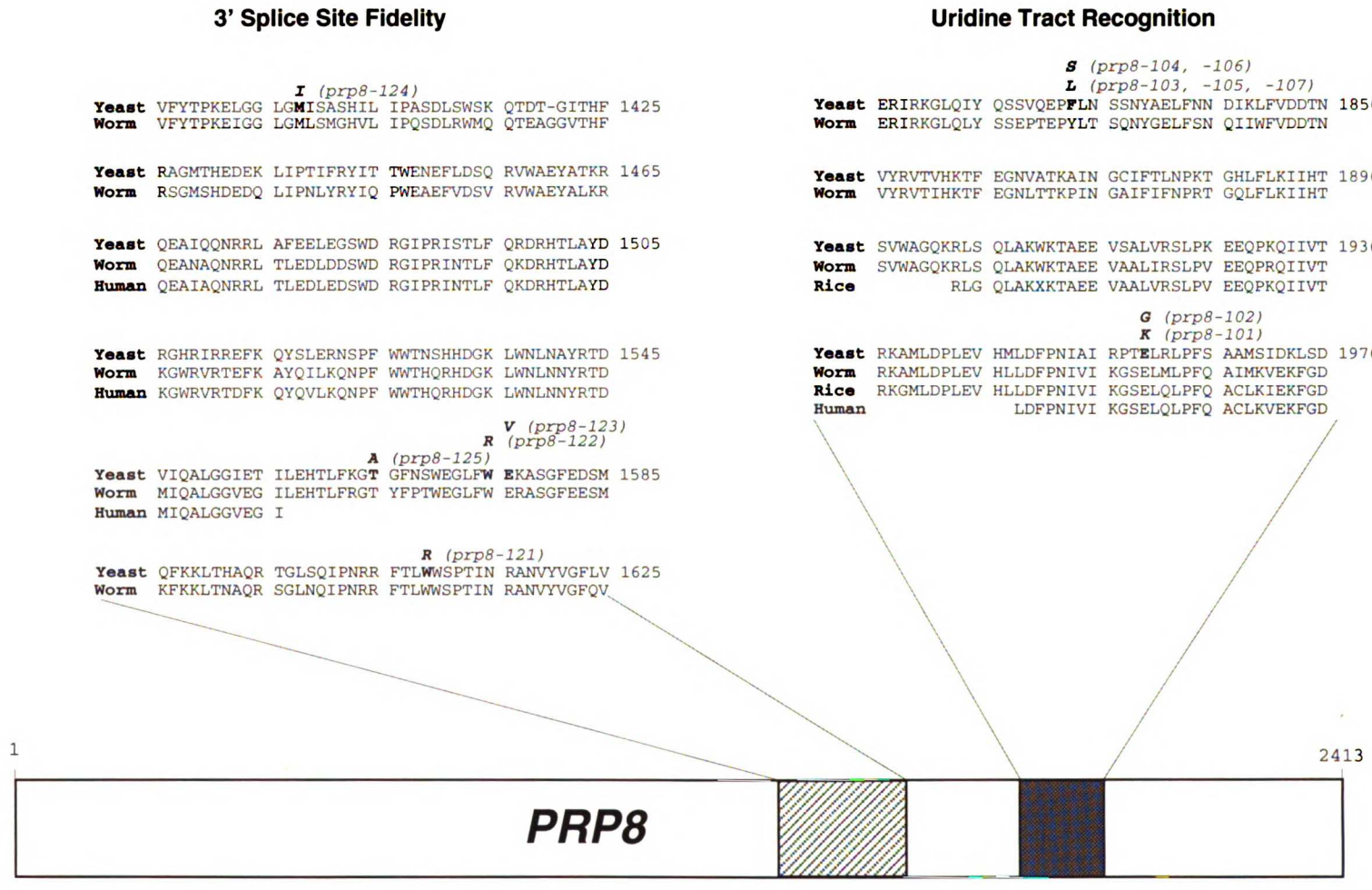

3' Splice Site Fidelity
I (prp8-124)
Yeast VFYTPKELGG LGMISASHIL IPASDLSWSK QTDT-GITHF 1425
Worm VFYTPKEIGG LGMLSMGHVL IPQSDLRWMQ QTEAGGVTHF
Yeast RAGMTHEDEK LIPTIFRYIT TWENEFLDSQ RVWAEYATKR 1465
Worm RSGMSHDEDQ LIPNLYRYIQ PWEAEFVDSV RVWAEYALKR
Yeast QEAIQQNRRL AFEELEGSWD RGIPRISTLF QRDRHTLAYD 1505
Worm QEANAQNRRL TLEDLDDSWD RGIPRINTLF QKDRHTLAYD
Human QEAIAQNRRL TLEDLEDSWD RGIPRINTLF QKDRHTLAYD
Yeast RGHRIRREFK QYSLERNSPF WWTNSHHDGK LWNLNAYRTD 1545
Worm KGWRVRTEFK AYQILKQNPF WWTHQRHDGK LWNLNNYRTD
Human KGWRVRTDFK QYQVLKQNPF WWTHQRHDGK LWNLNNYRTD
V (prp8-123)
R (prp8-122)
A (prp8-125)
Yeast VIQALGGIET ILEHTLFKGT GFNSWEGLFW EKASGFEDSM 1585
Worm MIQALGGVEG ILEHTLFRGT YFPTWEGLFW ERASGFEESM
Human MIQALGGVEG I
R (prp8-121)
Yeast QFKKLTHAQR TGLSQIPNRR FTLWWSPTIN RANVYVGFLV 1625
Worm KFKKLTNAQR SGLNQIPNRR FTLWWSPTIN RANVYVGFQV
Uridine Tract Recognition
S (prp8-104, -106)
L (prp8-103, -105, -107)
Yeast ERIRKGLQIY QSSVQEPFLN SSNYAELFNN DIKLFVDDTN 1856
Worm ERIRKGLQLY SSEPTEPYLT SQNYGELFSN QIIWFVDDTN
Yeast VYRVTVHKTF EGNVATKAIN GCIFTLNPKT GHLFLKIIHT 1896
Worm VYRVTIHKTF EGNLTTKPIN GAIFIFNPRT GQLFLKIIHT
Yeast SVWAGQKRLS QLAKWKTAEE VSALVRSLPK EEQPKQIIVT 1936
Worm SVWAGQKRLS QLAKWKTAEE VAALIRSLPV EEQPRQIIVT
Rice RLG QLAKXKTAEE VAALVRSLPV EEQPKQIIVT
G (prp8-102)
K (prp8-101)
Yeast RKAMLDPLEV HMLDFPNIAI RPTELRLPFS AAMSIDKLSD 1976
Worm RKAMLDPLEV HLLDFPNIVI KGSELMLPFQ AIMKVEKFGD
Rice RKGMLDPLEV HLLDFPNIVI KGSELQLPFQ ACLKIEKFGD
Human LDFPNIVI KGSELQLPFQ ACLKVEKFGD
1
2413
PRP8

TABLE 1

Characterization of mutagenized *PRP8* library

Mutagenized Region	#Transformants	% Null	# Temp. Sensitive/50	% *prp8-101*-like (total # isolated)	% 3' splice site suppressors (total # isolated)
Gap A	~500	4	n.d.	n.d.	n.d.
+PCR A	~4000	18	n.d.	0 (0)	0 (0)
+PCR A +Mn^{++}	~4000	76	n.d.	0 (0)	0 (0)
Gap B	~800	4	0	n.d.	n.d.
+PCR B	~4000	34	0	0 (0)	.025 (1)
+PCR B +Mn^{++}	~4000	48	9	0 (0)	.025 (1)
Gap C	~800	4	0	n.d.	n.d.
+PCR C	~4000	40	2	.4 (16)	.2 (8)
+PCR C +Mn^{++}	~4000	64	1	.075 (3)	.1 (4)
Gap D	~800	8	0	n.d.	n.d.
+PCR D	~4000	32	3	.55 (22)	.13 (5)
+PCR D +Mn^{++}	~4000	62	5	.1 (4)	0 (0)

The first column indicates the region of the gene that was mutagenized (Gap A-D) and cotransformed alone or with PCR DNA (+PCR A-D) or with manganese mutagenized PCR DNA (+PCR +Mn^{++} A-D). The second column indicates the number of transformants screened and the third column indicates the percentage null alleles generated (scored as inviable on 5FOA-containing media). The fourth column indicates the number of temperature sensitive mutants obtained from 50 randomly chosen transformants. The fifth column indicates the percentage of mutants with a *prp8-101*-like phenotype (loss of uridine recognition) with the actual number isolated indicated in parentheses. The last column indicates the percentage of mutants that suppress 3' GAG (described in text) with the actual number isolated indicated in parentheses. n.d. indicates that this parameter was not assayed.

TABLE 2

Phenotypes of new *prp8-101*-like alleles

Alelle	RNA	Primer Extension Analysis		Copper
		MP/MD	Loss of Uridine Preference	Resistance
PRP8	+T PyDOWN	26	n.a.	.05 mM
prp8-101	+T PyDOWN	2.0	13X	.18 mM
prp8-102	+T PyDOWN	4.9	5.3X	.15 mM
prp8-103	+T PyDOWN	3.6	7.2X	.15 mM +/-
prp8-104	+T PyDOWN	3.1	8.4X	.18 mM
prp8-105	+T PyDOWN	5.2	5.0X	.13 mM +/-
prp8-106	+T PyDOWN	2.2	12X	.15 mM
prp8-107	+T PyDOWN	5.3	4.9X	.15 mM
PRP8	+A WT	2.3	n.a.	n.d.
prp8-101	+A WT	7.4	3.2X	n.d.
prp8-102	+A WT	6.5	2.8X	n.d.
prp8-103	+A WT	3.7	1.6X	n.d.
prp8-104	+A WT	5.1	2.2X	n.d.
prp8-105	+A WT	7.9	3.4X	n.d.
prp8-106	+A WT	7.6	3.3X	n.d.
prp8-107	+A WT	3.9	1.7X	n.d.

The first column indicates the *PRP8* allele assayed. The second column indicates the reporter construct that was utilized. The third column shows quantitation of primer extension data as the ratio of branchsite proximal (MP) to branchsite distal (MD) mRNA from each of the competing 3' splice sites from Figure 3. The fourth column represents the amount of uridine preference lost in each mutant. For +T PyDOWN this value is (MP/MD)wild type/(MP/MD)mutant. For +A WT this value is (MP/MD)mutant/(MP/MD)wild type. The last column represents maximum copper resistance of each strain with the branchsite distal 3' splices site (uridine poor 3' splice site for +T PyDOWN) in frame with the *CUP1* sequence. n.a. indicates not applicable. n.d. indicates that the parameter was not measured.

TABLE 3

3' splice site suppression with construct set I

		Primer Extension Analysis		Copper Growth	
Strain	RNA	M/LI	Fold Suppression	Copper Growth	Fold Suppression
PRP8	3' GAG	.66	n.a.	.1 mM	n.a.
prp8-121	3' GAG	3.5	5.3X	.25 mM	2.5X
prp8-122	3' GAG	3.9	5.9X	.25 mM	2.5X
prp8-123	**3' GAG**	8.5	13X	.75 mM	7.5X
prp8-124	**3' GAG**	5.8	8.8X	.75 mM	7.5X
prp8-125	**3' GAG**	10.1	15X	.75 mM	7.5X
PRP8	3' UUG	.64	n.a.	.05 mM	n.a.
prp8-121	**3' UUG**	7.1	11X	.75 mM	15X
prp8-122	**3' UUG**	5.3	8.3X	.75 mM +/-	15X
prp8-123	3' UUG	1.3	2.0X	.25 mM	5X
prp8-124	3' UUG	2.1	3.3X	.25 mM	5X
prp8-125	3' UUG	.78	1.2X	.18 mM	4X

The first column indicates the strain utilized. The second column indicates the 3' splice site reporter from construct set I that each strain contained. In bold is each *PRP8* allele and the 3' splice site mutant against which it was selected. The third and fourth column represent quantitation of primer extension experiments measuring the efficiency of the second step as the ratio of mature message (M) to lariat intermediate (LI). Fold suppression is the fold increase in M/LI in the mutant versus wild type strains. The fifth and sixth columns represent copper resistance of each strain and the fold increase in copper resistance for each mutant versus wild type. n.a. indicates not applicable.

TABLE 4

Specificity of 3' splice site suppressors

Strain	RNA	Copper Growth/ β-gal units	Strain	RNA	Copper Growth/ β-gal units
PRP8	G5A Nl	.25 mM +/-	*PRP8*	+T PyDown	.05 mM
prp8-121	G5A Nl	< .05 mM	*prp8-121*	+T PyDown	.05 mM
prp8-122	G5A Nl	< .05 mM	*prp8-122*	+T PyDown	.05 mM
prp8-123	G5A Nl	.25 mM +/-	*prp8-123*	+T PyDown	.05 mM
prp8-124	G5A Nl	.1 mM +/-	*prp8-124*	+T PyDown	.05 mM
prp8-125	G5A Nl	.25 mM	*prp8-125*	+T PyDown	.05 mM
PRP8	G5A Ab	.25 mM +/-	*PRP8*	rp51a 3' UGG	1.0
prp8-121	G5A Ab	< .05 mM	*prp8-121*	rp51a 3' UGG	3.5
prp8-122	G5A Ab	< .05 mM	*prp8-122*	rp51a 3' UGG	10
prp8-123	G5A Ab	.25 mM	*prp8-123*	rp51a 3' UGG	1.2
prp8-124	G5A Ab	.05 mM	*prp8-124*	rp51a 3' UGG	1.9
prp8-125	G5A Ab	.25 mM +/-	*prp8-125*	rp51a 3' UGG	15
PRP8	A259C	.18 mM			
prp8-121	A259C	.1 mM			
prp8-122	A259C	.1 mM			
prp8-123	A259C	.1 mM			
prp8-124	A259C	.18 mM			
prp8-125	A259C	.1 mM			

The first column indicates the strain utilized and the second column indicates the reporter constructs. G5A Nl and G5A Ab indicate normal and aberrant splice sites are in frame with *CUP1* in this 5' splice site mutant. A259C is a branch site mutant and +T PyDOWN is desribed in Figure 2 and the legend for Table 2. rp51a 3' UGG is described in the text. Column 3 represents maximum copper resistance for each mutant or, in the case of rp51a 3' UGG, β-galactosidase activity in arbitrary units. +/- indicates weak growth at that concentration of copper.

TABLE 5

3' splice site suppression with construct set II:

3' UAG, AAG, GAG, UGG

		Primer Extension Analysis		Copper Growth	
Strain	RNA	M/LI	Fold Suppression	Copper Growth	Fold Suppression
PRP8	3' UAG (WT)	n.d.	n.a.	2 mM	n.a.
prp8-121	3' UAG (WT)	n.d.	n.a.	1.5 mM	none
prp8-122	3' UAG (WT)	n.d.	n.a.	1.5 mM	none
prp8-123	3' UAG (WT)	n.d.	n.a.	1.5 mM	none
prp8-124	3' UAG (WT)	n.d.	n.a.	1.5 mM	none
prp8-125	3' UAG (WT)	n.d.	n.a.	1.5 mM	none
PRP8	3' AAG	n.d.	n.a.	1.5 mM	n.a.
prp8-121	3' AAG	n.d.	n.a.	1.5 mM	none
prp8-122	3' AAG	n.d.	n.a.	1.5 mM	none
prp8-123	3' AAG	n.d.	n.a.	1.5 mM	none
prp8-124	3' AAG	n.d.	n.a.	1.5 mM	none
prp8-125	3' AAG	n.d.	n.a.	1.5 mM	none
PRP8	3' GAG	1.8	n.a.	.18 mM +/-	n.a.
prp8-121	3' GAG	18	10X	.75 mM +/-	4X
prp8-122	3' GAG	9.9	6X	.5 mM +/-	3X
prp8-123	3' GAG	9.6	5X	1 mM	6X
prp8-124	3' GAG	18	10X	1 mM	6X
prp8-125	3' GAG	26	14X	1 mM	6X
PRP8	3' UGG	1.1	n.a.	.075 mM +/-	n.a.
prp8-121	3' UGG	11	10X	.5 mM +/-	7X
prp8-122	3' UGG	4.1	4X	.18 mM +/-	2X
prp8-123	3' UGG	8.7	8X	.75 mM +/-	10X
prp8-124	3' UGG	7.6	7X	.25 mM	3X
prp8-125	3' UGG	6.7	6X	.25 mM	3X

See legend for Table 3.

TABLE 6

3' splice site suppression with construct set II: 3' UUG, UAU, UAC

			Cryptic 3' Splice Site Usage		Authentic 3' Splice Site Usage	
Strain	RNA	$M_{Wt}/M_{cryptic}$	M_{crypt}/LI	Fold Suppression[c]	M_{Wt}/LI	Fold Suppression[c]
PRP8	3' UUG	.78	.018	n.a.	.014	n.a.
prp8-121	3' UUG	1.0	.15	8X	.15	11X
prp8-122	3' UUG	.44	.11	6X	.048	3X
prp8-123	3' UUG	1.2	.10	6X	.13	9X
prp8-124	3' UUG	1.1	.081	5X	.087	6X
prp8-125	3' UUG	.98	.096	5X	.094	7X
PRP8	3' UAU	1.4	.035	n.a.	.049	n.a.
prp8-121	3' UAU	1.7	.25	7X	.43	9X
prp8-122	3' UAU	.91	.33	9X	.30	6X
prp8-123	3' UAU	2.4	.32	9X	.77	16X
prp8-124	3' UAU	2.3	.13	4X	.30	6X
prp8-125	3' UAU	2.1	.22	6X	.45	9X
PRP8	3' UAC	.21	.21	n.a.	.047	n.a.
prp8-121	3' UAC	.47	.66	3X	.31	7X
prp8-122	3' UAC	.30	.67	3X	.21	4X
prp8-123	3' UAC	.82	.36	2X	.30	6X
prp8-124	3' UAC	.69	.19	none	.13	3X
prp8-125	3' UAC	.49	.21	none	.099	2X

The first and second columns indicate the strain and 3' splice site reporter construct from set II. The third column is the ratio of wild type (M_{Wt}) to cryptic (M_{Crypt}) mature message produced. The fourth and fifth columns show quantitation of the efficiency of cryptic 3' splice site usage measured by the ratio of (M_{Crypt}) to lariat intermediate (LI) and compared to wild type to give fold suppression. The sixth and seventh columns show similar data for the wild type 3' splice site. n.a. indicates not applicable.

TABLE 7

Mutations in PRP8 identified in this study

PRP8 Allele	Mutation(s)	# Isolates
prp8-101	**E1960K**	2
prp8-102	**E1960G**	3
prp8-103	**F1834L**, V1946A	1
prp8-104	**F1834S**, R1922G	1
prp8-105	**F1834L**, F1880S	1
prp8-106	**F1834S**	1
prp8-107	**F1834L**	2
prp8-121	**W1609R**, N1618D	1
prp8-122	**W1575R**	1
prp8-123	**E1576V**, S1705C[a], N1730Y[a]	1
prp8-124	**M1399I**	1
prp8-125	**T1565A**, N1721Y, V1752A	1

The first column indicates the allele designation for each *PRP8* mutation. The second column indicates the mutations found by sequencing regions of each mutant. The alteration responsible for the selected phenotype with each allele is in bold. The number of isolates containing each mutation is shown in column three.

[a] These two mutations give weak 3' splice site suppression when present without the E1576V alteration. However, the E1576V change alone is sufficient to give the full level of 3' splice suppression seen in the original *prp8-123* isolate.

EPILOGUE

The Second Catalytic Step of pre-mRNA Splicing

INTRODUCTION

A ubiquitous feature of eukaryotes is the presence of intervening sequences which interrupt the coding regions of genes. Nuclear pre-mRNA splicing is the process by which these intervening sequences (introns) in messenger RNAs are precisely removed and the functional coding sequences (exons) ligated. Splicing proceeds via two transesterification reactions. In the first reaction (step one), the 2' hydroxyl group of an intron adenosine residue attacks the 5' splice site phosphodiester bond producing a branched lariat intermediate structure and a free 5' exon. In the second reaction (step two), the 3' hydroxyl group of the 5' exon attacks the 3' splice site phosphodiester bond producing ligated exons and an excised lariat intron (Figure 1).

The two catalytic steps of splicing are carried out by a large, ribonucleoprotein machine, the spliceosome. The main components of the spliceosome are five small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) which are packaged as ribonucleoprotein particles (snRNPs). In addition, there are multiple non-snRNP proteins that interact transiently with the spliceosome. The discovery of snRNAs as essential components of the spliceosome and the fact that group II self-splicing introns use a similar two-step chemical pathway of intron excision has led to the hypothesis that pre-mRNA splicing and group II splicing share a common evolutionary origin. Indeed, it is now generally believed that the catalytic mechanism of pre-mRNA splicing will be largely, if not solely, RNA based.

Unlike group II introns, which are "hard wired" with the proper RNA structures required for catalysis, the spliceosome is assembled *de novo* onto each intron that is removed. It then carries out the catalytic reactions and is presumably disassembled. One role that spliceosomal proteins are thought to play is mediating the elaborate series of RNA conformational changes required during spliceosome assembly and catalysis (Madhani & Guthrie, 1994a). The dynamic nature of the spliceosome cycle is summarized briefly below. For comprehensive reviews and citations, the reader is referred elsewhere (Green, 1991; Guthrie, 1991; Rymond & Rosbash, 1992; Moore et al., 1993).

The first step in assembly is binding of U1 snRNP to the 5' splice site via base-pairing between U1 snRNA and the intron. While this initial step is ATP-independent, each additional step in the assembly reaction requires ATP. Following U1, U2 snRNP binds the intron by base-pairing to sequences that flank the branchsite adenosine. The adenosine residue itself is not base paired but bulged out of the U2-branchsite helix allowing its utilization as a nucleophile in the first catalytic reaction (Query et al., 1994). After U2 snRNP binding, U4, U5 and U6 snRNPs enter the spliceosome, apparently as a triple snRNP particle. In this particle, U4 and U6 snRNAs are extensively base paired to each other. After binding of the U4/U5/U6 triple snRNP, the U4-U6 base-pairing interaction is disrupted and U4 is destabilized from the spliceosome. U6 is then able to isomerize into a base-pairing interaction with U2 that is mutually exclusive with its U4 interaction. The U1-5' splice site base-pairing interaction is also disrupted prior to the first catalytic step and is replaced by an interaction between the 5' splice site and U5 and U6 snRNAs. The network of snRNA-snRNA and snRNA-intron interactions that is thus formed is thought to be the structural basis for the active site in the first catalytic step (Figure 2)(Madhani & Guthrie, 1994a; Nilsen, 1994; Sun & Manley, 1995).

The focus of this review is events that occur subsequent to spliceosome assembly and the first catalytic step, namely 3' splice site selection and the second catalytic step. Genetic, biochemical, and chemical data reviewed below demonstrate that the second catalytic step of splicing involves a unique set of RNA-RNA and protein-RNA interactions that is distinct from those occurring during spliceosome assembly and the first catalytic step. These data are discussed with an emphasis on their possible mechanistic implications for 3' splice site selection and catalysis.

CHEMICAL CONSIDERATIONS

Recently, powerful technological advances in constructing and manipulating RNAs have made it possible to ask questions about the nature of specific reactive groups (Moore

& Sharp, 1992). Modification of the 2' hydroxyl group in the ribose ring of either the first or last intron G residue appears to have little effect on the first catalytic step of the reaction. In contrast, a 2'-OCH_3, but not a 2'-H, in the first intron G residue slows the rate of the second step more than ten fold. When these substituents are in the ribose ring of the last G residue, the rate of the second step is slowed 20- and 7-fold, respectively (Moore & Sharp, 1992). The magnitude of these effects suggests that the 2' hydroxyl groups are not critical for the chemical mechanism of splicing *per se*, but might be involved in positioning of the reactive groups or in a recognition event unrelated to catalysis.

A second type of experiment with site-specifically-modified RNAs involves the placement of stereochemical probes in the reactive 5' and 3' splice site phosphates. Inversion of the stereochemistry at phosphorothioates in both steps of splicing reveals that the chemical steps involve two in-line S_N2 displacement reactions (Maschhoff & Padgett, 1993; Moore & Sharp, 1993). The fact that both reactions are inhibited by the R_p but not S_p phosphorothioate stereoisomer would suggest that the second step of splicing does not proceed as a reversal of the first catalytic step, as in group I self splicing introns, but involves a significant conformational rearrangement resulting in a different active site for step two (McSwiggen & Cech, 1989; Rajagopal et al., 1989; Suh & Waring, 1992; Moore & Sharp, 1993). However, a general two-metal-ion mechanism has been proposed for group I, group II, and pre-mRNA splicing catalysis within a single active site (Steitz & Steitz, 1993). Even though the second step of splicing is a reversal of the first in this model, the mechanism can accommodate the stereochemical data from pre-mRNA splicing because a simple rotation around the scissile phosphate allows the pro Rp oxygen in step two to bind the catalytic magnesium ions in the same orientation as they are bound by the pro Rp oxygen in step one. Even if there are minimal structural rearrangements between the chemical steps of pre-mRNA splicing, there must be at least two alterations to the active site. (1) The branch lariat must be displaced, and (2) the 3' splice site phosphodiester bond must be positioned for attack by the hydroxyl group of the last residue in exon 1.

The nature of the chemistry for each step (e.g. 2'-5' versus 3'-5' bond formation) further mandates that the substrate binding sites for each reaction cannot be identical, though there could be some overlap (e.g. a single exon 1 binding site used for both steps). The data reviewed below do indeed suggest that there is a significant structural change between the two steps of splicing. However, it remains to be determined whether this rearrangement reflects the creation of a different active site for step two or the minimal remodeling required for use of a step one-like active site.

RNAs REQUIRED FOR THE SECOND CATALYTIC STEP

Pre-mRNA Sequences

Introns are defined by consensus sequences at the 5' splice site, branchsite, and 3' splice site. The 5' splice site consensus is **R/GU**AUGU in yeast (*Saccharomyces cerevisiae*) and G/**GU**RAGY in mammals. The branchsite consensus is UACUA**A**CA in yeast and YNYUR**A**CN in mammals. The 3' splice site consensus sequence is **YAG/** in yeast and **CAG/** in mammals. (/= cleavage site; R=purine; Y=pyrimidine; N= any nucleotide; the branch adenosine is underlined). In yeast these sequences are highly conserved among introns, whereas in mammals they are more divergent (data from Rymond & Rosbash, 1992). Nucleotides in bold have been demonstrated to be required for the second catalytic step of splicing (Jacquier et al., 1985; Newman et al., 1985; Parker & Guthrie, 1985; Reed & Maniatis, 1985; Ruskin & Green, 1985; Ruskin et al., 1985; Aebi et al., 1986; Fouser & Friesen, 1986; Hornig et al., 1986; Vijayraghavan et al., 1986; Fouser & Friesen, 1987; Freyer et al., 1987; Query et al., 1994). The effects of branchsite mutations on step two are particularly difficult to detect in mammalian systems due to cryptic branchsite activation and in yeast because they tend to destabilize the lariat intermediate (Padgett et al., 1985; Ruskin et al., 1985; Reed & Maniatis, 1988; Burgess & Guthrie, 1993b; Query et al., 1994) Therefore, lack of boldface does not necessarily indicate that a nucleotide has no role in the second catalytic step of the reaction.

Besides these consensus sequences, a polypyrimidine tract is usually found between the branchsite and 3' splice site. Although the branchsite and 3' splice site are generally close together (15-40 nucleotides), this spatial coupling is not obligatory (e.g. Schatz et al., 1986; Helfman & Ricci, 1989; Smith & Nadal-Ginard, 1989; Goux-Pelletan et al., 1990). The 5' splice site and branchsite sequences (and the pyrimidine tract in mammals) are required for both the first and second catalytic steps. The 3' splice site YAG motif is dispensable for the first step of splicing and is only absolutely required for the second catalytic step (Frendewey & Keller, 1985; Reed & Maniatis, 1985; Ruskin & Green, 1985; Rymond & Rosbash, 1985; Rymond et al., 1987; Reed, 1989). In many mammalian and S. pombe introns, the YAG does appear necessary for the first step, but this requirement can be alleviated by strengthening the pyrimidine tract or by improving base-pairing with U1 (see below) (Reed, 1989; Reich et al., 1992) Thus, the 3' splice site can be recognized at least twice during splicing, with at least one recognition event occuring prior to the first step (Zhuang & Weiner, 1990).

It is not immediately obvious why the 5' splice site and branchsite sequences should be necessary for the second catalytic step of splicing after fulfilling their requirement in the first catalytic step. However, the dual role that these elements play may serve several purposes. First, although not obligatory, the close proximity of the branchsite to the 3' splice site may facilitate identification of a 3' splice site after the first catalytic step. For this mechanism to operate, the branch structure must be identified and used as a marker to search for a downstream 3' splice site. In mammalian *in vitro* splicing, there is evidence for a scanning mechanism that begins at the branch point (Smith et al., 1989; see below).

A second reason for participation of the 5' splice site and branchsite nucleotides in the second catalytic step is that they could serve to enhance the fidelity of splicing. After the first catalytic step, there may be inspection mechanisms that determine whether proper sequences were utilized to form the lariat intermediate. Incorrect intermediates could be prevented from continuing the reaction or be eliminated. Indeed, such a mehcanism has

been found recently in yeast. The spliceosomal ATPase Prp16 (see below) controls a pathway which degrades lariat intermediates formed at mutant branchsites (Burgess & Guthrie, 1993b).

A final reason for requiring the 5' splice site and branchsite in the second catalytic step is that the nucleotides in the branch lariat structure may be direct participants in the reaction. This requirement would provide a means of ensuring coupling between the two steps of splicing because 3' splice site cleavage and ligation would depend on proper 5' splice site cleavage and lariat formation. A highly insightful experiment has provided evidence that this is the case. It had been known for some time that mutations in first G of the intron allow the first step of splicing to proceed, albeit at a reduced rate. However, these mutations completely block the second step (Newman et al., 1985; Aebi et al., 1986; Fouser & Friesen, 1986; Vijayraghavan et al., 1986). Similarly, mutations in the last G residue also strongly block the second step of splicing (Vijayraghavan et al., 1986). Amazingly, when a G to A mutation in the first residue of the yeast actin intron is combined with a G to C mutation in the last intron residue, splicing is restored to relatively high levels (10% of wild type message) (Parker & Siliciano, 1993). A G to A change in the last position also functions to give mutual suppression in combination with the G to A change at position one, though to a lesser extent. The same interaction has also been demonstrated with the yeast rp51a intron (Chanfreau et al., 1994). The allele-specific nature of the suppression is strongly suggestive of a direct interaction between the first and last G residues. Thus, the 5' splice site G residue in the branch lariat structure provides a specific recognition element for the G in the YAG motif at the 3' splice site.

Similar results have now been obtained with the /A-C/ combination in a mammalian intron (Deirdre et al., 1995). Interestingly, it was also found that inosine, when substituted for guanosine in both the first and last intron positions, has little effect on the rate of the second catalytic step. This information, combined with possible non-canonical pairing schemes that accommodate I-I, G-G and A-C with similar geometries, allows a reasonable

guess as to the nature of the G-G interaction (Figure 3) (Tinoco, 1993; Deirdre et al., 1995). Direct proof of this interaction will require a more thorough substitution analysis of the relevant guanosine ring substituents or perhaps identification of a crosslink between the first and last intron residues. To date, no further evidence has been obtained for other intron-intron interactions between conserved nucleotides (Chanfreau et al., 1994; Ruis et al., 1994; Deirdre et al., 1995).

The YAG motif is the only 3' splice site sequence element that is absolutely required for the second catalytic step. Mutations in these three nucleotides inhibit the second step of splicing in the following order of increasing severity Py < A < G (Vijayraghavan et al., 1986; Fouser & Friesen, 1987; Parker & Siliciano, 1993; Chanfreau et al., 1994). Therefore, besides the aforementioned /G-G/ interaction, there must be other recognition interactions involving these nucleotides during the second step. Base-pairing of the 3' splice site AG dinucleotide with the 5' end of U1 snRNA has been shown to be required for the splicing of some *Schizosaccharomyces pombe* introns, but only for the first step of splicing (Reich et al., 1992). A test of the same base-pairing scheme in *S. cerevisiae* failed to turn up evidence for this interaction in either step of splicing (Seraphin & Kandels-Lewis, 1993; Umen & Guthrie, unpublished).

Besides the YAG trinucleotide, an upstream pyrimidine tract contributes to 3' splice site recognition during the second step of splicing in both mammals and yeast. In mammals, this requirement is easily obscured by the strong requirement for the pyrimidine tract in spliceosome assembly prior to the first catalytic step (Frendewey & Keller, 1985; Reed & Maniatis, 1985; Ruskin & Green, 1985). However, with the appropriate substrate, a role for the pyrimidine tract in the second catalytic step has been uncovered (Reed, 1989). In yeast, the pyrimidine tract is less conserved than in mammals and is restricted to mostly uridine residues. Analysis of the yeast intron database reveals a strong enrichment for uridine residues preceding the YAG motif in yeast 3' splice sites. One position in particular, -9 with respect to the G at the 3' cleavage site, is conserved as a

uridine in over 80% of yeast introns and could be considered as a 3' splice site consensus element (Parker & Patterson, 1987; Rymond & Rosbash, 1992). The functional role of the pyrimidine tract in yeast was demonstrated by making use of a sensitive 3' splice site competition assay. When two 3' splice sites are linked in cis, if one contains a uridine-rich tract it has a strong competitive advantage over the uridine-poor 3' splice site (Patterson & Guthrie, 1991).

In addition to the YAG motif and pyrimidine tract, there appear to be other elements that can contribute to 3' splice site recognition. One of these, discussed below, is an interaction with U5 snRNA and exon positions adjacent to the 3' splice site. There are also "context" effects that appear when the 3' splice site sequence is sub-optimal. For example, when the actin intron 3' splice site is mutated from UAG/ to UUG/, most of the splicing still occurs at this mutated cleavage site and not at a UUG sequence upstream in the intron, even though the upstream UUG is in what appears to be a favorable location. Furthermore, a second, cryptic 3' splice site with the sequence AUG/ is activated by the UUG/ mutation (Parker & Siliciano, 1993); this site is 5 nucleotides upstream of the utilized UUG/ 3' splice site. Why is this double mutant cryptic site chosen over the more upstream UUG sequence, and how is the spliceosome able to choose the "correct" UUG? Use of these 3' splice sites does not correlate with U5 snRNA base-pairing potential at the exon sequences (see below). Therefore, there must be additional information in the pre-mRNA that can specify possible splice sites. This information does not appear to be encoded in consensus sequence elements. Instead, it might be correlated with overall sequence content. Yeast intron sequences are more AU rich than exons (Parker & Patterson, 1987); therefore, the intron/exon borders could be crudely defined by the changes in sequence content flanking the splice junction. A striking example of a role for sequence context in splicing is found in plants. Plant introns can be visually identified by a very strong difference in AU content between introns and exons. Furthermore, AU richness (versus a specific sequence motif) has been shown to play a prominent functional

role in the splicing of plant introns (Goodall & Filipowicz, 1989). More work is necessary to determine whether such a mechanism is utilized in other organisms.

U2 and U6 snRNAs

Only three of the spliceosomal snRNAs, U2, U5 and U6, appear to be required for the catalytic steps of splicing. U1 and U4 are both destablized from the spliceosome prior to the first catalytic step, and splicing catalysis has been observed in their absence (Yean & Lin, 1991; Crispino et al., 1994; Tarn & Steitz, 1994). As mentioned above, during spliceosome assembly, U2 and U6 snRNAs isomerize into a base-pairing interaction that is required for the first catalytic step of splicing. This interaction consists of two short helices (Ia and Ib) interrupted by a two nucleotide bulge in U2 (See figure 2)(Madhani & Guthrie, 1992). The accumulation of lariat intermediates that is seen when helix I is disrupted indicates that it is required for the second catalytic step. In mammals, there are additional base-pairing interactions formed between U2 and U6, but their role in step two has not been determined (Datta & Weiner, 1991; Wu & Manley, 1991; Sun & Manley, 1995).

Two clusters of individual nucleotides in U2 and U6 snRNAs are necessary specifically for the second catalytic step. The first set lies just upstream of helix I in the phylogenetically conserved ACAGAG sequence in U6. Alterations in the first four bases of this sequence cause a block to the first step of splicing whereas mutations in the last two (boldface) bases block the second catalytic step *in vitro* and *in vivo* (Fabrizio & Abelson, 1990; Madhani et al., 1990; Vankan et al., 1992; Datta & Weiner, 1993; Yu et al., 1993; Wolff et al., 1994). The other set of nucleotides that are required for the second step of splicing are in the bulge region between helices Ia and Ib (Fabrizio & Abelson, 1990; Madhani et al., 1990; Madhani & Guthrie, 1992; McPheeters & Abelson, 1992; Madhani & Guthrie, 1994b; Wolff et al., 1994). These include U6 residues C58 and A59 (from yeast) opposite the bulge and their U2 base-pairing partners U23 and G26. Alterations in U2 A27

cause a modest decrease in the second catalytic step. Some discrepancies between data from yeast and mammalian systems suggests that the step two nucleotides in U2 and U6 are not always rate limiting (Vankan et al., 1992; Datta & Weiner, 1993; Wolff et al., 1994).

While part of the second step splicing block caused by mutations around the bulge region of helix I is due to disruption of helices Ia and Ib, there is a phenotypic asymmetry resulting from alterations in U6 at position 59 versus its U2 pairing partner U23 that cannot be explained by disruption of helix Ib. *In vivo*, mutations in yeast U6 position 59 lead to temperature sensitive growth or lethality, whereas position 23 mutations in U2 cause no growth phenotypes (Madhani & Guthrie, 1992). This asymmetry suggests that A59 in U6 plays an additional role in the second catalytic step besides base-pairing with U2. There is also evidence suggesting that the bulged structure itself in helix I is important for the second step. Although the U2 nucleotides in the bulge can be substituted with other bases without greatly affecting splicing efficiency (McPheeters & Abelson, 1992; Madhani & Guthrie, 1994c), altering the spacing of the bulge by insertion or deletion of nucleotides is highly deleterious (Madhani & Guthrie, 1994c; Wolff et al., 1994).

Interestingly, substitution of U2 position 25 in the bulge affects 3' splice site utilization. Changes at this position slightly decrease the *in vivo* efficiency of step two for wild type substrates, and an A to G change in U2 25 can suppress the effects of alterations at the 3' splice junction AG dinucleotide. The suppression is not allele specific, but could be taken to suggest that the U2 mutation is altering the active site to decrease its selectivity (Madhani & Guthrie, 1994c). There is a mutation in yeast U6, G52U, that results in similar non-specific suppression of 3' splice site AG alterations (Lesser & Guthrie, 1993). Moreover, these two residues in U2 and U6 are implicated in a possible tertiary interaction involving a non-canonical G-A pairing. Evidence for the tertiary interaction comes from a covariation seen between these nucleotides after randomization and selection for functional variants of U2 and U6 *in vivo* (Madhani & Guthrie, 1994c). As with U6 A59/U2 U23,

there is a phenotypic asymmetry observed between U2 position 25 mutations and U6 position 52 mutations, indicating that participation of U6 G52 in the proposed tertiary interaction is secondary to a more critical role for this nucleotide. Nevertheless, the tertiary interaction is capable of bringing together the two clusters of U2 and U6 residues that are required for the second catalytic step and is supported by crosslinking data (see below).

Intriguingly, in mammalian extracts, a crosslink can be isolated between the nucleotide adjacent to the tertiary interaction in mammalian U6 (underlined adenosine residue in the U6 ACAG**A**G motif) and position two in the intron (underlined U in G/G**U**RAGY at the 5' splice site) (Sontheimer & Steitz, 1993). Moreover, a different crosslink has recently been identified between yeast U2 residue U23 and the first nucleotide of exon 2 (A. Newman, personal communication). These crosslinks both occur in the lariat intermediate and suggest that all the nucleotides important for the second catalytic step in U2 and U6 can be simultaneously juxtaposed with critical substrate residues in the intron prior to catalysis (Figure 4).

A second type of analysis with U6 snRNA has involved substitution of phosphates with phosphorothioates. These experiments were performed by *in vitro* reconstitutions of U6 in both nematodes and yeast. The results of the two studies are highly congruent in their identification of Rp phosphorothioates that block each step of splicing. In particular, the phosphate between positions 58 and 59 in yeast U6 (positions 48 and 49 of *Ascaris lumbricoides*) is required specifically for the second step of splicing (Fabrizio & Abelson, 1992; Yu et al., 1995). Notably, this phosphate is positioned in the "kink" of the bulge region that is critical for the second catalytic step (see figure 4). An interesting possibility is that this phosphate is responsible for binding a catalytic magnesium ion in the active site. Though the inability to rescue the defect with manganese does not rule out this possibility (Fabrizio & Abelson, 1990; Yu et al., 1995), there is, as yet, no direct evidence for specific magnesium binding sites in the spliceosome. Identification of such sites is of obvious

importance since it is likely that magnesium ions participate in splicing catalysis (Piccirilli et al., 1993; Steitz & Steitz, 1993).

U5 snRNA

The role of U5 snRNA in either step of splicing has until recently seemed enigmatic. The finding that depletion of U5 can block the second step of splicing *in vivo* and *in vitro* suggests that U5 is not absolutely required for the first step (Patterson & Guthrie, 1987; Winkelmann et al., 1989). However, this interpretation runs counter to the standard view of spliceosome assembly and to other reports that demonstrate a stringent U5 snRNP requirement for spliceosome assembly and for the first catalytic step (Lamm et al., 1991; Seraphin et al., 1991; Segault et al., 1995). This discrepancy cannot be explained solely by the extent of U5 depletion in the respective experiments. That is, if any spliceosome that can carry out the first step of splicing contains U5 snRNP, then all spliceosomes that undergo the first catalytic step must contain U5 and therefore should be able to carry out the second step. One resolution of this paradox is possible if U5 is absolutely required for both steps of splicing but is exchangeable between spliceosomes and can dissociate after the first catalytic step. A more conventional explanation invokes the existence of exchangeable U5 snRNP proteins that are required for the second step of splicing. If the second step U5 proteins must compete for binding with other U5 proteins, then depletion of U5 might increase the concentration of free U5 proteins to the point where the second step proteins are outcompeted. Interestingly, there is one yeast U5 snRNP protein, Prp18, that behaves as an exchangeable factor and is required for the second catalytic step (see below) (Horowitz & Abelson, 1993b).

The most conserved portion of U5 snRNA is a 9 nucleotide loop at the top of a stem structure (Guthrie & Patterson, 1988; Frank et al., 1994). A breakthrough in deciphering the function of U5 has come through the isolation of loop mutants that can suppress splice site mutations or affect cryptic splice site usage in yeast. Careful genetic

analysis has shown that four nucleotides in the loop (in boldface; 5' $G_1C_2C_3\mathbf{U_4U_5U_6U_7A_8C_9}$ 3') can base pair with both exons adjacent to the 5' and 3' splice sites (Newman & Norman, 1991; Newman & Norman, 1992) (Figures 2 and 4). Activation of cryptic 5' splice sites can occur in yeast when the normal 5' splice site is mutated and the UU residues at loop positions 5 and 6 can base pair with positions -2 and -3 with respect to the cryptic cleavage site. Genetic experiments with the U5 loop in mammals have yielded similar results and possibly demonstrated an extension of the base-pairing to include an interaction between loop position 4 and position -1 of the first exon (Cortes et al., 1993). With respect to the second step of splicing, the effects of alterations in the 3' splice site AG dinucleotide can be partially suppressed when complementarity is created between positions 4 and 5 of the U5 loop (CU) and exon 2 positions +1 and +2 (Newman & Norman, 1992). It is notable that the two U5-exon interactions described are out of register with one another in terms of forming a continuously base paired structure.

Importantly, the interactions between the exons and the loop of U5 have been demonstrated for a wild type intron which lacks the potential to base pair with the loop nucleotides (Wyatt et al., 1992; Sontheimer & Steitz, 1993). 4-thiouridine placed at position -1 in exon 1 of a mammalian intron crosslinks to U5 loop nucleotides 4 and 5 prior to step one, and this U5-pre-mRNA crosslink can be "chased" through both steps of splicing, indicating its functional relevance (Sontheimer & Steitz, 1993). Both exon 1 and exon 2 crosslinks are in the proper register with respect to the genetic interactions described for each exon (Sontheimer & Steitz, 1993). 4-thiouridine at exon 2 position +1 crosslinks to loop nucleotides 3 and 4 only after the first step of splicing. In summary, the loop nucleotides of U5 are in close proximity to the reactive groups that participate in the second catalytic step (See figure 4).

Despite the wealth of detailed information on the U5 loop, its specific function is still unclear. Since exon sequences are not well conserved, the general function of the U5 loop nucleotides cannot be to specify splice site location through base-pairing with exons.

It has been proposed that the function of the loop is to hold exon 1 after the first step of splicing and to help align the exons for the second step of splicing (Newman & Norman, 1992; Sontheimer & Steitz, 1993). Nonetheless, the limited and variable extent of this interaction suggests that other factors are required for its stabilization.

It is important to note that the above sections have focused on nucleotides and phosphates that are required only for the second catalytic step. It is highly likely that there is also some overlap with residues involved in the first catalytic step. Knowledge of which residues are specific to step one and step two and which are required for both catalytic steps will be invaluable for understanding the relationship between the step one and step two active sites. Even without this information, however, the interactions that have already been inferred from genetic and biochemical experiments allow an impressive number of constraints to be imposed on the architecture of the putative active sites. These constraints must be viewed with some caution, of course, since they are often predicated on the assumption that the RNA-RNA interactions described are occurring simultaneously. The nucleotides and RNA-RNA interactions required for the second catalytic step are summarized in Figure 4.

PROTEINS REQUIRED FOR THE SECOND CATALYTIC STEP

Just as there are RNA sequences required specifically for the second step of splicing, there is a set of proteins that is also required for step two (see Table 1). In addition, there is at least one protein, Prp8, that is required for both steps of splicing. Because more information is available on these proteins in yeast, the yeast data will be discussed first, followed by a summary of mammalian proteins involved in the second step of splicing.

PRP16

PRP16 was first identified from a mutant, *prp16-1,* that suppresses the splicing defect of an A to C change at the intron branch nucleotide (Couto et al., 1987; Burgess et al., 1990). Subsequently, it was identified as a temperature sensitive mutant, *prp23-1*, blocked at the second catalytic step (Vijayraghavan et al., 1989). *In vitro*, Prp16 is required only for the second step (Schwer & Guthrie, 1991). The sequence of the gene reveals it to be a member of the DEXH family of putative RNA helicases (Burgess et al., 1990; Schmid & Linder, 1992). Indeed, Prp16 has been shown to be an RNA dependent ATPase, although no RNA helicase activity has been found associated with the purified protein (Schwer & Guthrie, 1991). *PRP16* shares sequence homology with two other splicing factors, *PRP2* and *PRP22*. All three belong to the DEAH subgroup and are tripartite in structure with a unique N terminal domain followed by homologous "helicase" and C terminal domains (Chen & Lin, 1990). Prp2 is required just prior to the first catalytic step, and Prp22 is required for the release of mature message from spliceosomes (Company et al., 1991; Kim & Lin, 1993).

Prp16 functions by binding to spliceosomes after lariat intermediate formation; it then hydrolyzes ATP and exits (Schwer & Guthrie, 1991). Part of its spliceosomal binding site may be the 3' splice site, since Prp16 can be specifically crosslinked to this region of the intron (Umen & Guthrie, 1995c). Moreover, ATP hydrolysis by Prp16 results in a conformational change that leads to protection of the 3' splice site from oligonucleotide-directed RNase H cleavage (Schwer & Guthrie, 1992a). This conformational change is correlated with 3' splice site crosslinking of two other proteins, Slu7 and Prp8 (see below) (Umen & Guthrie, 1995c). The conformational change could reflect a function for Prp16 in bringing a potential 3' splice site into the spliceosomal active site or in altering the spliceosome for the second catalytic step.

If Prp16 is involved in remodeling the spliceosome for step two, then it would be expected to interact with spliceosomal snRNAs required for this step. Interestingly, U2

and U6 snRNAs are among the strongest stimulators of the RNA-dependent ATPase activity of Prp16 *in vitro* (Yan Wang, personal communication). Moreover, a genetic interaction with U6 snRNA was found by isolating U6 suppressors of dominant negative, cold sensitive *PRP16* alleles (Madhani & Guthrie, 1994b). These dominant negative alleles are thought to block a reaction after spliceosomal binding of Prp16, e.g. release of the protein after ATP hydrolysis (Schwer & Guthrie, 1992b). In this model, the U6 suppressors would act to partially disrupt the Prp16 binding site, thus allowing the protein to exit the spliceosome after carrying out its function. The strongest of these suppressors are single nucleotide deletions in U6 just upstream of the ACAGAG motif in a heptanucleotide sequence AAACAAU (nucleotides 40-46; see Figure 4) (Madhani & Guthrie, 1994b). As these nucleotides are not known to be critical for either step of splicing, they may form a secondary or redundant binding site for Prp16. Disruption of the site would then provide enough destablilization for suppression but not enough to interfere significantly with wild type functions of U6 or Prp16.

How do these data fit with the original identification of *prp16-1* as a suppressor of an intron branchsite mutation? The *prp16-1* allele was found to contain a single substitution near a conserved NTP binding motif in the helicase domain (Burgess et al., 1990). Not surprisingly, this allele causes a severe defect in RNA-dependent ATPase activity (Schwer & Guthrie, 1992b). Subsequent work in which additional branchsite suppressor alleles were isolated and analyzed demonstrated that these alleles all map to the helicase domain and cause a decrease in RNA-dependent ATPase activity. It was also determined that the suppressors do not function by accelerating the rate of the second catalytic step with mutant branch lariats. Instead, the suppressor alleles inhibit the degradation of these aberrant lariats (Burgess & Guthrie, 1993b). The mechanism of suppression can be explained if there is a rate limiting step (LI->LI*) for productive splicing of mutant lariat intermediates, and this rate-limiting step is in kinetic competition with ATP hydrolysis by Prp16 (Burgess & Guthrie, 1993b; Burgess & Guthrie, 1993a).

The rate-limiting step might be the aforementioned conformational change that results in 3' splice site protection (Schwer & Guthrie, 1992a). This idea is appealing given that the branch lariat structure itself could be part of a binding site for the 3' splice site and would thus contribute energetically to the conformational change. ATP hydrolysis by Prp16 would lock the lariat-bound 3' splice site into place and allow splicing to proceed. However, mutant branchsites would impair 3' splice site binding and prevent the LI->LI* transition from taking place before ATP hydrolysis, thus dooming the lariat intermediate to degradation. In this model, some spliceosomal binding of the 3' splice site can take place independently of ATP hydrolysis by Prp16. This facet of the model is supported by the finding of 3' splice site crosslinks to a spliceosomal protein, Prp8, in the absence of Prp16 (see below) (Umen & Guthrie, 1995c).

One important and unresolved issue is how many rounds of ATP hydrolysis are required to complete the Prp16-dependent (and other ATP-dependent) steps of splicing. Whether there are multiple rounds of hydrolysis or a single event would have a large impact on how we think about the mechanism by which Prp16 functions. For example, 3' splice site selection may involve several rounds of ATP hydrolysis accompanied by the binding and release of one or more potential 3' splice sites. ATP hydrolysis by Prp16 (or another ATPase) could prevent the binding reaction from reaching equilibrium and thereby enhance the fidelity of splice site selection in a manner akin to kinetic proofreading during translation (Hopfield, 1974; Ninio, 1975).

SLU7

SLU7, like *PRP16*, encodes an essential function and is required for the second step of splicing, both *in vivo* and *in vitro* (Frank & Guthrie, 1992; Frank et al., 1992; Ansari & Schwer, 1995; Jones et al., 1995). This gene was isolated as *slu7-1* in a screen for mutations that are synthetically lethal with loop mutations in U5 snRNA. The *SLU7* sequence contains a short stretch of homology to a family of retroviral capsid proteins

(Frank & Guthrie, 1992). This sequence motif ($CX_2CX_4HX_4C$), termed a "zinc knuckle", is implicated in RNA packaging by retroviruses (Rein, 1994). Thus, Slu7 is predicted to be an RNA binding protein. Although Slu7 does not appear to be stably associated with U5 snRNP (D. Frank and C. G., unpublished data), a clue to a possible RNA binding site came from experiments using splicing substrates with tandem 3' splice sites competing in cis. The *slu7-1* allele causes a selective defect in utilization of 3' splice sites that are greater than ~12 nucleotides downstream of the branchsite (Frank & Guthrie, 1992). Therefore, Slu7 is likely to be involved in utilization of 3' splice sites that are distal to the branchsite. Supporting the idea of a direct role for Slu7 in 3' splice site selection is the finding of a Slu7-3' splice site crosslink during the second step of splicing (Umen & Guthrie, 1995c). It has not been determined whether complete absence of the protein differentially affects distal 3' splice sites or whether this is a special property of the *slu7-1* allele. Despite predictions from sequence homology, mutations in the zinc knuckle motif of *SLU7* cause only a modest decrease in cell growth and splicing and do not appear to affect 3' splice site selection or 3' splice site crosslinking significantly (Frank & Guthrie, 1992; Umen & Guthrie, unpublished data).

In vitro experiments with Slu7 have demonstrated that this protein does not require ATP to carry out its function in the second step of splicing (Ansari & Schwer, 1995; Jones et al., 1995). This finding suggests that Slu7 acts after the Prp16-dependent step, which does require ATP. A more direct test of this idea utilized purified spliceosomes lacking both Prp16 and Slu7. Adding back Prp16, Slu7, or ATP in a specified order confirmed that Prp16 functions prior to Slu7 and that Slu7 functions in the absence of ATP. Additional fractionation experiments suggest the existence of a novel second step factor, SSF1, that is required in addition to Prp16 and Slu7. The biochemical properties of SSF1 indicate that it is probably not identical to Prp17 or Prp18 (see below) (Ansari & Schwer, 1995).

Consistent with its function after ATP hydrolysis by Prp16, Slu7 crosslinks most strongly to the 3' splice site after the Prp16-dependent step of the reaction. This crosslinking also depends on the functions of other known second step splicing factors, Prp17 and Prp18. Thus, Slu7 interacts with the 3' splice site at a time very close to the second catalytic event (Umen & Guthrie, 1995c). One of its functions could be to mediate or stabilize binding of the 3' splice site to the spliceosome before or after Prp16 exits. 3' splice sites that are distal to the branchsite might be particularly sensitive to loss of Slu7 function since they are expected to have a slower "on" rate for spliceosome binding due to their greater distance from the binding site.

Prp17 and Prp18

Both *PRP17* and *PRP18* were identified from temperature sensitive mutants, *prp17-1* and *prp18-1*, that specifically block the second step of splicing (Vijayraghavan et al., 1989; Vijayraghavan & Abelson, 1990). *slu4-1*, which is synthetically lethal with a loop mutation in U5 snRNA, is also a mutant allele of *PRP17* (Frank et al., 1992). The *PRP18* sequence reveals no homologies to known proteins. *PRP17* contains four copies of the WD motif, a segment of approximately 40 amino acids that is thought to mediate protein-protein interactions (Neer et al., 1994). Interestingly, the genes that encode Prp17 and Prp18 are not essential, and as expected, absence of either protein causes only a partial block to the second step of splicing *in vitro* (Horowitz & Abelson, 1993b; Jones et al., 1995). Prp18 can be found associated with the U4/U5/U6 triple snRNP, probably as a U5 snRNP component (Horowitz & Abelson, 1993b). Because depletion of Prp18 from extracts does not result in a first step splicing block, co-immunodepletion of U4/U5/U6 must only be limited. Thus, Prp18 associates with these snRNPs relatively weakly (i.e. in an exchangeable manner) or with only a subset of triple snRNPs (Horowitz & Abelson, 1993a).

Like Prp16 and Slu7, Prp17 and Prp18 have been functionally ordered with respect to the ATP requirement during the second step of splicing. Prp17 acts before or concomitant with an ATP-dependent reaction whereas the function of Prp18 is ATP-independent (Horowitz & Abelson, 1993a; Jones et al., 1995). Thus, a strict ordering would place the Prp16/Prp17-dependent functions prior to the Slu7/Prp18-dependent functions (see Figure 6). Perhaps the non-essential partner in each of these pairs facilitates the function of the essential factor. The observation that overexpression of Prp16 suppresses a *PRP17* mutation and that overexpression of Slu7 suppresses a *PRP18* mutation support this idea (Jones et al., 1995). Using information on ATP requirements, protection and crosslinking, some of the events that take place during the second step of splicing can be tentatively ordered with respect to protein factor requirements (Figure 6).

The formal ordering of these factors described above may actually prove misleading for understanding their functions. In the strictest sense, the "upstream" factors should not be affected by the "downstream" ones. This relationship is violated, however, by the finding that release of Prp16 from the 3' splice site partly depends on a "downstream" protein, Slu7 (Umen & Guthrie, 1995c). Genetic experiments, in particular, suggest close functional connections that are independent of ATP requirements (Figure 5). Alleles of *SLU7*, *PRP17* and *PRP18* are all synthetically lethal with each other and with loop mutations in U5. Alleles of *PRP16* are synthetically lethal with alleles of *PRP17* and *SLU7* but not *PRP18* (Frank et al., 1992). Overexpression of *PRP16* can suppress mutations in *SLU7* as well as in *PRP17* (Jones et al., 1995). Finally, alleles of all four genes, *PRP16*, *PRP17*, *PRP18* and *SLU7*, are synthetically lethal with an allele of *PRP8*, *prp8-101* (See below; Figure 5) (Umen & Guthrie, 1995c). Extensive genetic interactions such as these are often indicative of physical interactions. To date, the best evidence of direct associations are between U5 snRNP and the proteins Prp18 and Prp8 (Lossky et al., 1987; Horowitz & Abelson, 1993b). Preliminary results suggest that Prp8 directly contacts U5 snRNA (C. Collins, personal communication.)

Prp8

Unlike the previous four proteins, Prp8 was first identified from a temperature sensitive mutant, *prp8-1*, that blocks splicing prior to the first catalytic step of the reaction *in vivo* and *in vitro* (Jackson et al., 1988; Brown & Beggs, 1992). Removal of the protein by genetic depletion also inhibits the first step of splicing. The consequence of removing Prp8 is destabilization of the U4/U5/U6 triple snRNP which, in turn, blocks spliceosome assembly (Brown & Beggs, 1992). Immunoprecipitations indicate that the protein is present in U5 snRNP, U4/U5/U6 snRNP, and spliceosomes (Lossky et al., 1987; Whittaker et al., 1990; Brown & Beggs, 1992). Anti-Prp8 antibodies were shown to cross-react with a mammalian U5 snRNP protein of similar size (220 kd), giving a first indication of the remarkable evolutionary conservation of *PRP8* (Anderson et al., 1989; Pinto & Steitz, 1989). Subsequent sequence comparisons revealed that the yeast and *C. elegans* homologs are approximately 68% identical and 80% similar over most of the entire length of both genes. The cloned regions of *PRP8* homologues from other species show similar levels of identity and similarity (Hodges et al., 1995).

Though the conservation and large size of Prp8 indicate that it likely plays critical roles in the splicing reaction, its sequence reveals no conserved motifs in common with other proteins that give a clue to its function. Crosslinking experiments in yeast and mammals have demonstrated that Prp8 or its mammalian homologue p220 directly contacts the pre-mRNA, lariat intermediate, and excised lariat intron (Garcia-Blanco et al., 1990; Whittaker & Beggs, 1991; Teigelkamp et al., 1995b). Subsequently, one contact point was found at position -2 of the first exon during mammalian *in vitro* splicing. This interaction takes place prior to the first catalytic step (Wyatt et al., 1992).

A specific functional role for Prp8 was first identified in 3' splice site selection. A novel allele, *prp8-101*, was found to block recognition of the uridine tract preceding the 3' splice site. As expected from its impaired 3' splice site recognition phenotype, this allele

inhibits the second catalytic step of splicing. It also strongly exacerbates the phenotypes of mutations in the YAG motif at the 3' splice junction (Umen & Guthrie, 1995b). Crosslinking of Prp8 to the 3' splice site suggests that its role in mediating 3' splice site selection is direct (Teigelkamp et al., 1995a; Umen & Guthrie, 1995b).

Further crosslinking of Prp8 to the splicing substrate has identified contacts with the first exon in positions -1, -2 and -8 prior to the first step of splicing. After the first step, Prp8 contacts the branchsite, 3' splice site, and part of the second exon (Teigelkamp et al., 1995a). Thus, Prp8 is a good candidate for a protein that assists the U5 loop in binding to exon sequences (see above). Detailed kinetic analysis of Prp8-3' splice site crosslinking indicates that it occurs weakly prior to hydrolysis of ATP by Prp16 and strengthens afterwards. Strong crosslinking is also dependent upon the functions of Prp17, Prp18 and Slu7. Slu7-3' splice site crosslinking follows a similar profile to that of Prp8 (Umen & Guthrie, 1995c). Comparison of these crosslinking kinetics to those of Prp16 suggests that the 3' splice site is recognized in at least two distinct stages during the second catalytic step. The first stage is characterized by strong crosslinking of Prp16 and the second by strong crosslinking of Prp8 and Slu7. This two-stage binding regime may reflect a proofreading mechanism that ensures proper 3' splice site selection. The crosslinking kinetics are also consistent with the presence of Prp8 and Slu7 at or near the 3' splice site during the second catalytic step. This result, combined with the high degree of sequence conservation of Prp8 in particular, is suggestive of a role for this protein at the active site.

Support for the idea that Prp8 may be involved in utilization of the YAG motif at the 3' splice junction comes from a novel class of *PRP8* alleles that is distinct from *prp8-101*. In contrast to *prp8-101*, which excacerbates the phenotypes of YAG alterations at the 3' splice site, these new alleles suppress YAG alterations without affecting uridine tract recognition (Umen & Guthrie, 1995a). This suppression is highly reminiscent of that seen when presumptive active site RNA residues, yeast U6 G52 or U2 A25, are mutated (see

above). While these *PRP8* alleles can suppress a wide spectrum of YAG alterations, the complex pattern of allele preference displayed suggests a direct interaction with the YAG trinucleotide and/or the active site.

Mammalian Second Step Splicing Factors

Study of the second step in mammalian splicing extracts has lagged somewhat compared to yeast, possibly because of the difficulty involved in reproducibly blocking this step *in vitro* or the low relative abundance of second step proteins. Nonetheless, there has been progress recently in isolating such proteins in mammals.

Early biochemical experiments yielded two fractions that contained essential second step activities. These were termed SF3 and SF4a (Krainer & Maniatis, 1985). SF3, a heat-labile activity, was later shown to function at or prior to an ATP-dependent step, making it analogous to Prp16 and Prp17 (Sawa et al., 1988). No further characterization of SF3 or SF4a has been reported. A fraction termed Ia, which also contains an essential second step activity, was independently isolated (Perkins et al., 1986). A more recent fractionation experiment has identified at least one other factor that appears to be distinct from SF3. The purest preparation of this fraction, the flowthrough from a CM-sepharose column (CMFT), was shown to act only at the second step of splicing and to function after the requirement for ATP. Thus, this activity might be analogous to Slu7 or Prp18 (Lindsey et al., 1995).

An alternative approach for identifying putative second step factors in mammals has been to purify large quantities of spliceosomes that are blocked for step two. This block is achieved by using a mutant intron that is missing the AG dinucleotide at the 3' splice site. The proteins associated with this substrate are expected to be enriched for second step factors (Gozani et al., 1994). Indeed, at least one mammalian second step protein, PSF (see below), is present in these spliceosomes along with fourteen novel proteins. It is not yet known at which steps of splicing the novel proteins act.

To date, only a single identified mammalian protein, PSF, has been shown to play a role in step two. PSF (PTB associated splicing factor) is a pyrimidine tract binding protein that is found complexed with another pyrimidine tract binding protein, PTB (Garcia-Blanco et al., 1989; Patton et al., 1993). PTB does not appear essential for splicing *in vitro*. Although originally characterized as a first step splicing factor, depletion of PSF also results in lariat intermediate accumulation. The second step block can be complemented with purified recombinant PSF (Gozani et al., 1994). The ATP requirement of PSF during the second step has not been reported.

A different line of experimentation has yielded the interesting conclusion that protein dephosphorylation also plays a role in the second step of splicing. It was found that okadaic acid, an inhibitor of type 2A (and to a lesser extent type 1) protein phosphatases, can specifically block the second step of splicing (Mermoud et al., 1992; Tazi et al., 1992). This result correlates with the finding that ATPγS also inhibits the second step, presumably by rendering a target protein phosphatase resistant (Tazi et al., 1992). Thus, it appears that one or more proteins must be dephosphorylated prior to the second catalytic step. These two inhibitors should prove useful for developing assays to identify the relevant phophatase and its substrate(s).

MECHANISMS OF 3' SPLICE SITE SELECTION

Some aspects of the mechanism of 3' splice site selection in yeast and mammals are worth comparison. It should be noted that for both systems, part of 3' splice site choice is dictated by the site of branch lariat formation during the first catalytic step. Here we will focus only on the mechanism of 3' splice site localization after lariat intermediate formation.

Using oligonucleotide-directed RNAse H cleavage as an assay, 3' splice site protection has been probed in both systems. In yeast, the 3' splice site is accessible to cleavage after the first catalytic step but becomes protected in an ATP- and Prp16-dependent reaction (see above). This protection is correlated with strong crosslinking of

Slu7 and Prp8 to the 3' splice site. In mammals, the 3' splice site is also accessible to cleavage after the first catalytic step but can be converted to a protected form with the addition of two fractions necessary for step two. However, this protection occurs in the absence of ATP. Subsequent addition of ATP to these spliceosomes is sufficient to allow the second step to proceed (Sawa & Shimura, 1991). The difference in ATP requirements may be due to the nature of the assay systems. The yeast experiments were in Prp16-depleted extracts and utilized a mutant 3' splice site, whereas the mammalian experiments were performed with spliceosomes immobilized on a solid support. In both systems, the 3' splice site undergoes at least one conformational change, from unprotected to protected, prior to the second catalytic step. As suggested above, the protection might arise from direct interaction of the 3' splice site with its spliceosomal binding site. The 3' splice site protection patterns seen in yeast and mammals are shown in Figure 6.

A difficult issue has been the mechanism of 3' splice site localization in yeast and mammals after the first catalytic step. As mentioned above, there is evidence for a 5' to 3' scanning mechanism in mammals that chooses the first AG dinucleotide downstream of the branchsite. The scanning "rule" is broken only in exceptional circumstances where the first AG is either very close to the branchsite (and probably sterically occluded) or directly adjacent to a downstream AG (Smith et al., 1989; Smith et al., 1993). In yeast, despite an inherent preference for branchsite proximal 3' splice sites, several lines of evidence argue against a simple scanning mechanism (Langford et al., 1984; Patterson & Guthrie, 1991). First, even in relatively poor sequence context (i.e. preceded by purines, not pyrimidines), a downstream AG can be somewhat competitive with an upstream AG (Patterson & Guthrie, 1991). Second, interposing secondary structures can be accommodated by the yeast splicing machinery (Halfter & Gallwitz, 1988; Deshler & Rossi, 1991). Finally, the sequence context of a downstream splice site in yeast can affect utilization of an upstream competitor (Patterson & Guthrie, 1991). Even if a "leaky" scanning mechanism were

invoked, it would not explain the ability of a downstream splice site to be utilized at the expense of an upstream competitor.

How can these observations be resolved? First, it should be noted that besides the difference in organism, the yeast data are based on *in vivo* observations whereas the mammalian data are from *in vitro* experiments. A recent attempt to look at competition between 3' splice sites in yeast extracts yielded the result that only branchsite proximal AG dinucleotides were utilized (Teigelkamp et al., 1995a). Thus, *in vitro* 3' splice site selection may involve different dynamics than the *in vivo* reaction. Competition between AGs has been observed in mammalian extracts in apparent violation of the above mentioned scanning rules (Reed, 1989). This experiment utilized a different splicing substrate than that used to infer the scanning mechanism. Therefore, the particular sequence context of the substrate or the splicing extract preparation may affect whether or not scanning takes place. In both yeast and mammals, branchsite proximal 3' splice sites are favored, suggesting that if scanning is not used, then a diffusion-collision mechanism is employed. That is, 3' splice sites that are closer to the branchsite would have a higher rate of interaction with the spliceosome and would be used more frequently than distal 3' splice sites. In summary, the two systems may utilize both scanning and diffusioncollision mechanisms for 3' splice site localization.

STEP TWO IN pre-mRNA VERSUS GROUP II SPLICING

Much attention is now being paid to the similarities between group II splicing and pre-mRNA splicing. Viewed on one level, these similarities are so striking as to leave no doubt that the two share a common origin. However, starting with the bias that specific details form a more informative basis of comparison, the situation becomes somewhat ambiguous. This ambiguity is particularly apparent for the step two-specific RNAs in each system.

Step two interactions in group II introns

The structure of group II introns can represented as a wheel with six domains of secondary structure (I-VI) forming radial spokes. Domain V is the most conserved in structure and sequence and lies just upstream of the bulged adenosine-containing branch site helix (domain VI), followed by the 3' splice site (Michel et al., 1989). The 3' splice site AY/ dinucleotide is recognized and aligned with the 5' exon by three known interactions. The first is with an exon binding site structure termed EBS1. EBS1 is in a loop atop a stem structure and forms six or seven continuous base pairs with exon 1 starting from the 5' cleavage site. Next to the paired 5' cleavage site is a position in EBS1 that base pairs with the first position in exon 2, acting as a guide to juxtapose the 3' exon precisely with the 5' exon in helical continuity (Jacquier & Michel, 1987; Jacquier & Jacquesson-Breuleux, 1991). The second interaction involved in 3' splice site recognition is a base pair between an intron nucleotide and the pyrimidine at the 3' splice site (γ-γ') (Jacquier & Michel, 1990; Jacquier & Jacquesson-Breuleux, 1991). The third is an interaction between the first intron residue, a G, and the penultimate residue, an A (Chanfreau & Jacquier, 1993).

As expected from its conservation and from mutational analysis, many of the functionally important residues in group II introns are found in domain V (Jarrell et al., 1988; Michel et al., 1989; Koch et al., 1992; Chanfreau & Jacquier, 1994; Boulanger et al., 1995; Peebles et al., 1995). This domain of the intron forms a helix-bulge-helix structure similar to spliceosomal helices Ia and Ib in the U2-U6 structure (See figure 7). Modification interference experiments identified several nucleotides and non-bridging phosphate oxygens in domain V that are critical for splicing (Chanfreau & Jacquier, 1994). Notably, there were no alterations that differentially affected step two. Outside of domain V, domain VI and the lariat structure also play roles in the second catalytic step (Schmelzer & Muller, 1987; Jacquier & Jacquesson-Breuleux, 1991; Koch et al., 1992). Domain VI

forms a tertiary interaction with domain II (η-η'), and probably with domain V (Dib-Hajj et al., 1993; Chanfreau & Jacquier, 1995).

Spliceosomal Versus Group II Second Step RNAs

There are some obvious comparisons between spliceosomal and group II RNAs required for the second step. The first is between EBSI and the conserved loop of U5 snRNA, both of which interact with the exons adjacent to the splice sites. Though these interactions are superficially similar, there are several important differences. Most apparent are the lack of base pairing potential between U5 and exon sequences. Furthermore, the U5 interaction with the exons is not in register for continuous basepairing. The extent of the interaction with exon 1 is also quite different in the two systems, being much larger in EBS I. While spliceosomal proteins or other RNAs may act to stabilize U5-exon interactions, on its own merits, the loop of U5 appears to be functionally quite distinct from EBSI. These differences do not rule out a common origin, but the functional divergence neutralizes any argument for or against this view.

A similar situation holds for the interaction between the intron ends that is required for step two in both systems. In the case of pre-mRNA splicing this interaction involves a G-G pairing at the first and last positions. In the case of group two introns it is a G-A pairing between the first and penultimate positions. The most likely pairing schemes based on suppression data do not share the same geometry (Chanfreau & Jacquier, 1993; Parker & Siliciano, 1993; Deirdre et al., 1995)(Figure 3). Therefore, the superficial similarity between these interactions does not generate a particularly illuminating comparison. It could be argued (as above) that interaction between intron ends is a useful strategy for coupling the two steps of splicing. Thus, the differences between the two systems would be as easily attributable to convergent evolution as to a common evolutionary origin (Weiner, 1993).

Perhaps the most important comparison between the two systems involves U2 and U6 snRNA structures and domain V, both of which are thought to be active site components. The original analogy between the two placed the bulge nucleotides in identical positions (Madhani & Guthrie, 1992) (see Figure 7, top). In this register, helices Ia and Ib of U6 correspond to the upper and lower helices of domain V. Ideally, alterations in the nucleotides of domain V would give similar phenotypes to their U6 counterparts. In particular, there should be nucleotides around the bulge region of domain V that differentially affect step two. As mentioned above, there is, to date, no evidence of any nucleotide, phosphate, or structure in domain V being required specifically for the second step of splicing. Although the bulge structure has not been extensively mutated, alterations seem to affect binding of domain V to the rest of the ribozyme prior to the first catalytic step (Chanfreau & Jacquier, 1994; Peebles et al., 1995).

A second analogy has been made between domain V and U2/U6 that places the conserved trinucleotide, AGC, which forms helix Ib in U2/U6, in register with a highly conserved AGC trinucleotide at the base of domain V (Boulanger et al., 1995; Peebles et al., 1995; Yu et al., 1995) (See figure 7, bottom). While lacking the simple geometrical similarities of the original model, this new register presents some striking parallels between group II and spliceosomal RNAs. The first is that mutations in the AGC of domain V block splicing as they do in U6. Furthermore, the asymmetry on either side of the helix is conserved. Mutation of the pairing partners for the AGC in domain V, like the analogous U2 mutations in helix Ib, has a much less severe effect on splicing than altering the AGC nucleotides themselves (Madhani & Guthrie, 1992; Chanfreau & Jacquier, 1994; Boulanger et al., 1995; Peebles et al., 1995). Finally, there is a striking similarity in the location of non-bridging phosphate oxygens required for the first step of splicing in both systems (Fabrizio & Abelson, 1992; Chanfreau & Jacquier, 1994; Yu et al., 1995). The first lies between the A and G in the trinucleotide and the second lies in the bulge of domain V, which is analogous to a bulge in an intermolecular stem loop of U6 in the revised

model. This new model raises the question of how to think about the step two specific residues in the U2/U6 structure since they now have no obvious analogs in domain V.

The above observations clearly do not rule out the possibility of a fundamental similarity between the mechanism of group II splicing and pre-mRNA splicing. Given the chemical differences in the two steps of splicing, it is somewhat surprising not to find some conserved nucleotides or structures in domain V that contribute specifically to this step of the reaction. It is possible that the mutations which affect step two in group II introns are not rate limiting or also block splicing at an earlier step. Even for pre-mRNA splicing, step two effects in U6 can be variable depending on the assay system utilized (Vankan et al., 1992; Datta & Weiner, 1993; Wolff et al., 1994).

Whether or not the similarities between group II introns and pre-mRNA splicing share a common origin or are the product of chemical determinism (Weiner, 1993), the possibility of either of these relationships has been a strong and synergistic driving force for progress in both fields over the past several years.

CONCLUSIONS

A growing amount of information is now available on proteins and RNAs that contribute to the second step of splicing. Besides their predicted role in modulating RNA rearrangements, second step proteins also mediate splice site selection and may have a role in forming the spliceosomal active site. Our focus must now turn to questions regarding details of how these proteins interact with each other, and with RNAs required for the second step. For example do the second step proteins interact physically to form a complex? Do Prp16 and Prp8 bind active site snRNAs? If so, what are the binding kinetics? With the answers to these questions, comes the the exciting prospect of integrating proteins into the ever more intricate network of RNA-RNA interactions that is formed during the second catalytic step.

Even without protein-based information, characterization of spliceosomal RNAs is bringing us tantalizingly close to a possible RNA-based active site model for step two (Figure 4). Yet, one largely missing piece of information from this network is the set of interactions responsible for binding the 3' splice site itself. The data available are consistent with YAG recognition taking place within the step two active site (as opposed to a binding site distinct from the site of catalysis). As already evidenced by the aforementioned G-G interaction, YAG recognition may be complicated, involving multiple non-canonical interactions between the 3' splice site and active site RNAs. However, determining the nature of this binding site could be especially rewarding. The additional spacial constraints that this information would add to current models may allow us to start thinking on a new level about the three dimensional architecture of the active site, and about possible catalytic mechanisms.

Acknowledgments

We are grateful to Cathy Collins, Hiten Madhani and Jon Staley for critical comments and discussions regarding this manuscript. J. G. U. was supported by a NIH training grant, NSF pre-doctoral fellowship, and NIH research grant GM21119 to C.G. C.G. is an American Cancer Society Research Professor of Molecular Genetics.

Figure 1. The two chemical steps of pre-mRNA splicing.

Exons are depicted as shaded rectangles. The yeast consensus sequences are shown with nucleotides required for the second catalytic step in bold. Y represents a pyrimidine (C or U) and p represents a reactive phosphate.

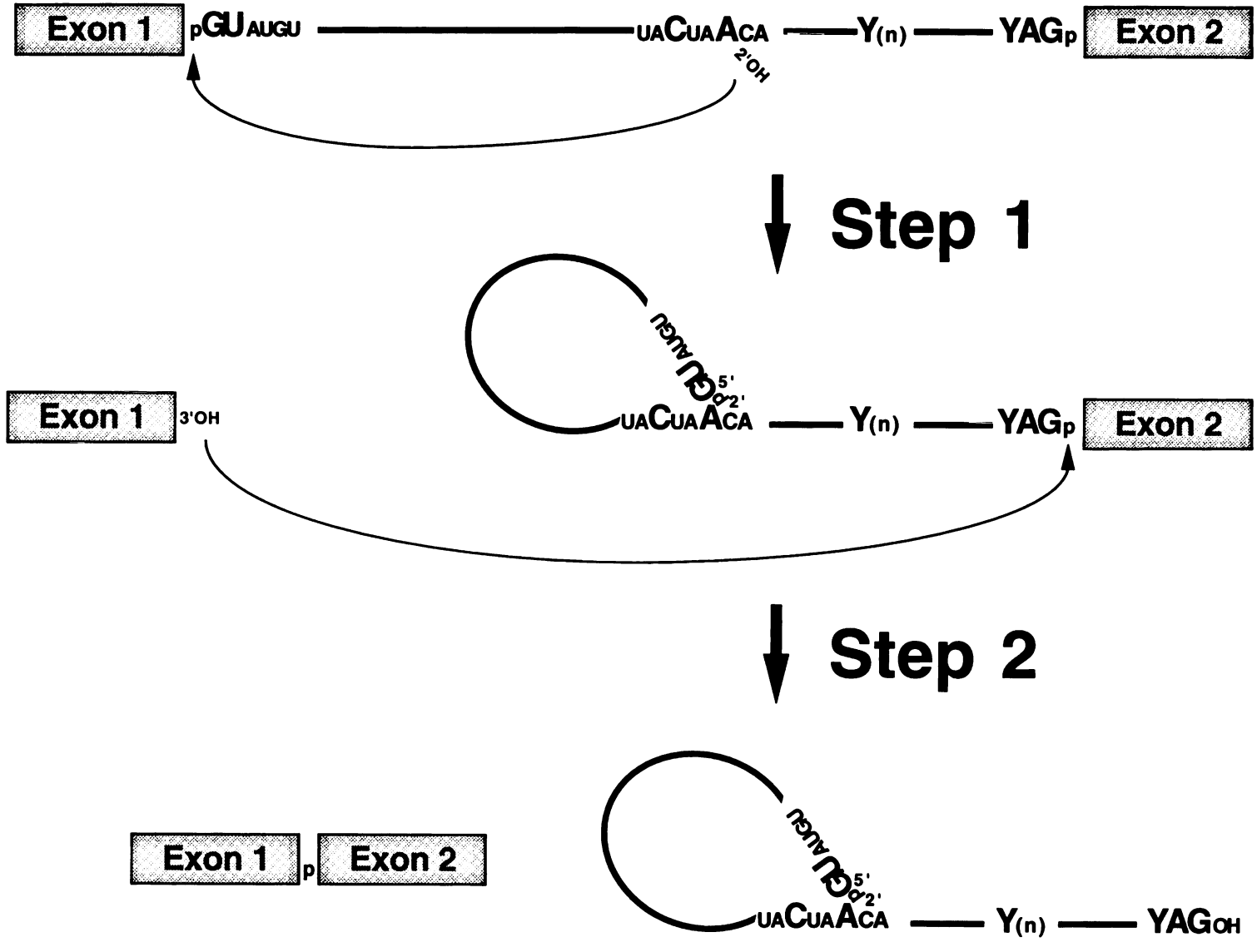

Exon 1
pGUAUGU
UACUAACA
2'OH
Y(n)
YAGp
Exon 2
Step 1
Exon 1
3'OH
pGUAUGU
5'
2'
UACUAACA
Y(n)
YAGp
Exon 2
Step 2
Exon 1
p
Exon 2
pGUAUGU
5'
2'
UACUAACA
Y(n)
YAGOH

Figure 2. A network of RNA interactions prior to the first catalytic step. The numbering scheme for *S. cerevisiae* snRNAs is used. Depicted are the U2-branchsite helix, the U6-5' splice site helix, U2-U6 helix I and the U5-exon 1 interaction. Exons are dark rectangles and the path of the intron is represented by the lightly shaded line. Residues that participate in the second catalytic step are highlighted in bold. Adapted from Madhani & Guthrie (1994a).

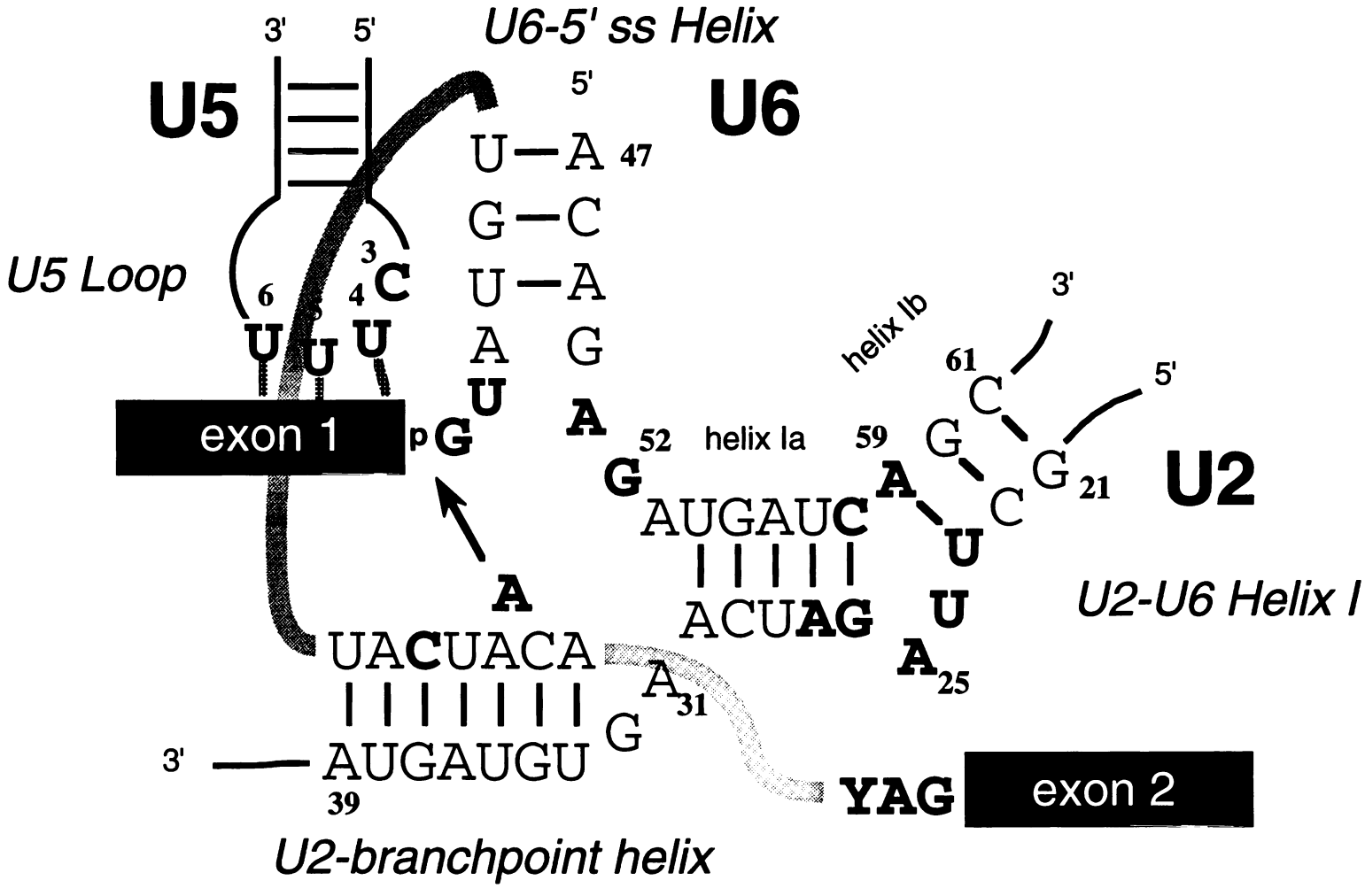
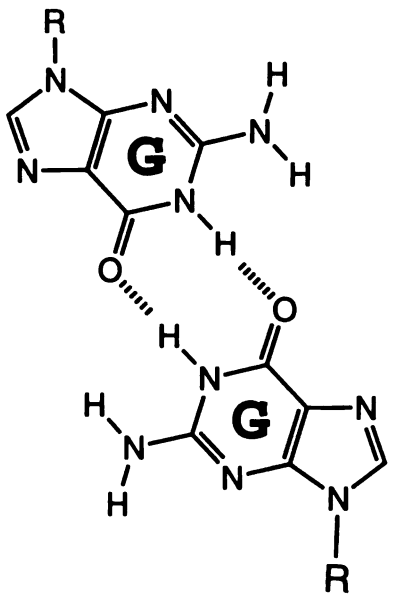
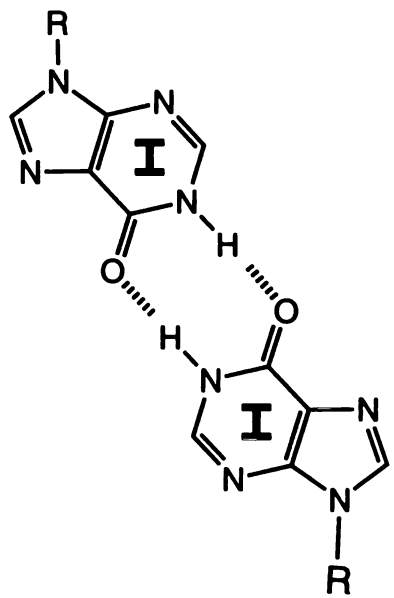

U6-5' ss Helix
3'
5'
U5
5'
U6
U—A 47
G—C
U—A
A G
U
A
52
G
U5 Loop
3
C
6
4
U
U
U
exon 1
pG
helix Ib
3'
61
C
5'
helix Ia
59
G
G
21
U2
A
C
AUGAUC
U
U2-U6 Helix I
A
ACUAG
U
A
25
UACUACA
A
31
G
3'
AUGAUGU
39
YAG
exon 2
U2-branchpoint helix

Figure 3. Non-Watson-Crick interaction between the first and last intron residues. (A) *G-G* hydrogen bonding thought to occur in wild type introns. (B) The wild type interaction is not disrupted when guanosine *G* is replaced with inosine *I*. (C) The adeonsine-cytosine *A-C* configuration is accomodated with some backbone distortion (note position of R-groups in (A) and (B) versus (C). Adapted from Tinoco (1993).

A.


R
N
N
G
H
N
H
N
H
O
O
H
N
G
H
N
H
N
N
N
R

B.


R
N
N
I
N
H
O
O
H
N
I
N
N
N
R

C.

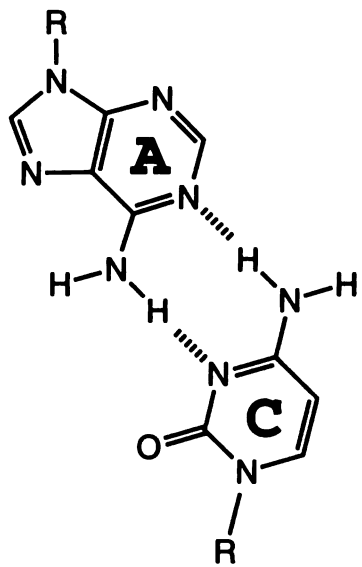

R
N
N
A
N
H
N
H
H
N
H
N
N
C
O
N
R

Figure 4. A network of RNA interactions prior to the second catalytic step. The numbering scheme for *S. cerevisiae* snRNAs is used. Depicted are the U2-branchsite helix, U6-5' splice site helix, U2-U6 helix I and U5-exon 1/exon 2 interactions after lariat intermediate formation. Exons are dark rectangles and the intron is represented by a lightly shaded thick line. The nucleotides and phosphates that participate in step two are shown in bold. The U6 heptanucleotide sequence that interacts genetically with *PRP16* is also shown. The 2'-5' A-G bond at the lariat branch is depicted as a thick dark line. The non-canonical G-G interaction between the first and last intron residues is depicted as a coarsely striped line, and the tertiary interaction between U2 A25 and U6 G52 is represented as a finely striped line. Crosslinks between U6 A51-intron +2 or U2 U23-exon two +1 are depicted as dark stippled lines. The phosphate whose non-bridging oxygen is important for step two is indicated by the shaded gray circle. Adapted from Madhani & Guthrie (1994b). See text for individual citations.

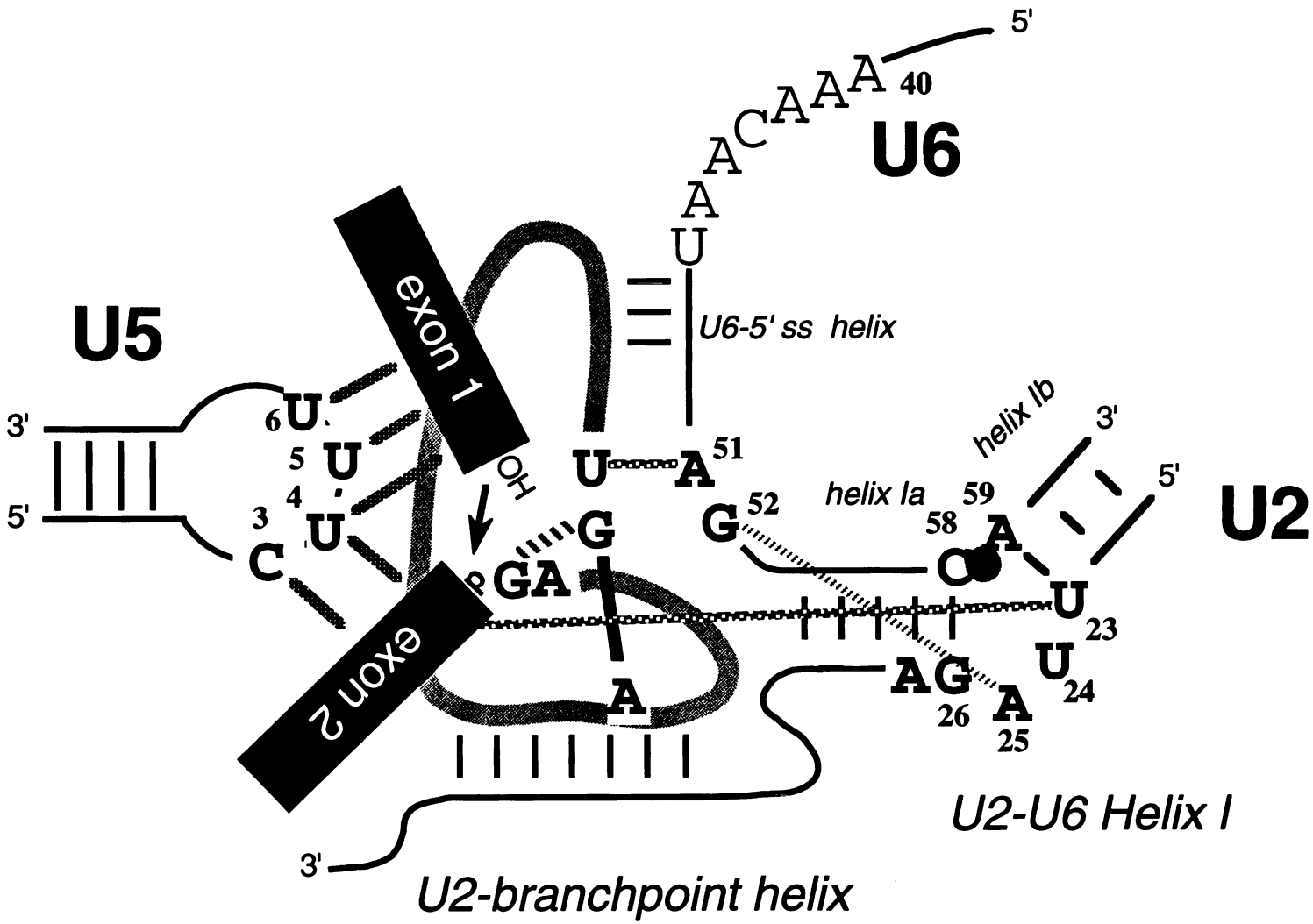

5'
A C A A A 40
U6
A
U
U6-5' ss helix
U5
exon 1
3'
5'
6 U
5 U
4 U
3
C
OH
U
A 51
G 52
helix Ib
3'
helix Ia
59
58 A
5'
U2
G
GA
U 23
U 24
AG
26 A
25
exon 2
A
3'
U2-U6 Helix I
U2-branchpoint helix

Figure 5. Genetic interactions between yeast second step splicing factors. Proteins are depicted as spheres and U5 snRNA as a sphere with a stem structure. Solid dark lines between spheres indicate that alleles of the genes that encode the splicing factors are synthetically lethal. Line thickness is used only to represent perspective. Lightly colored arrows indicate genetic suppression by overexpression. Overexpression of the factor at the beginning of the arrow results in partial suppression of the growth phenotype caused by a mutation in the factor at the arrow head. See text for citations.

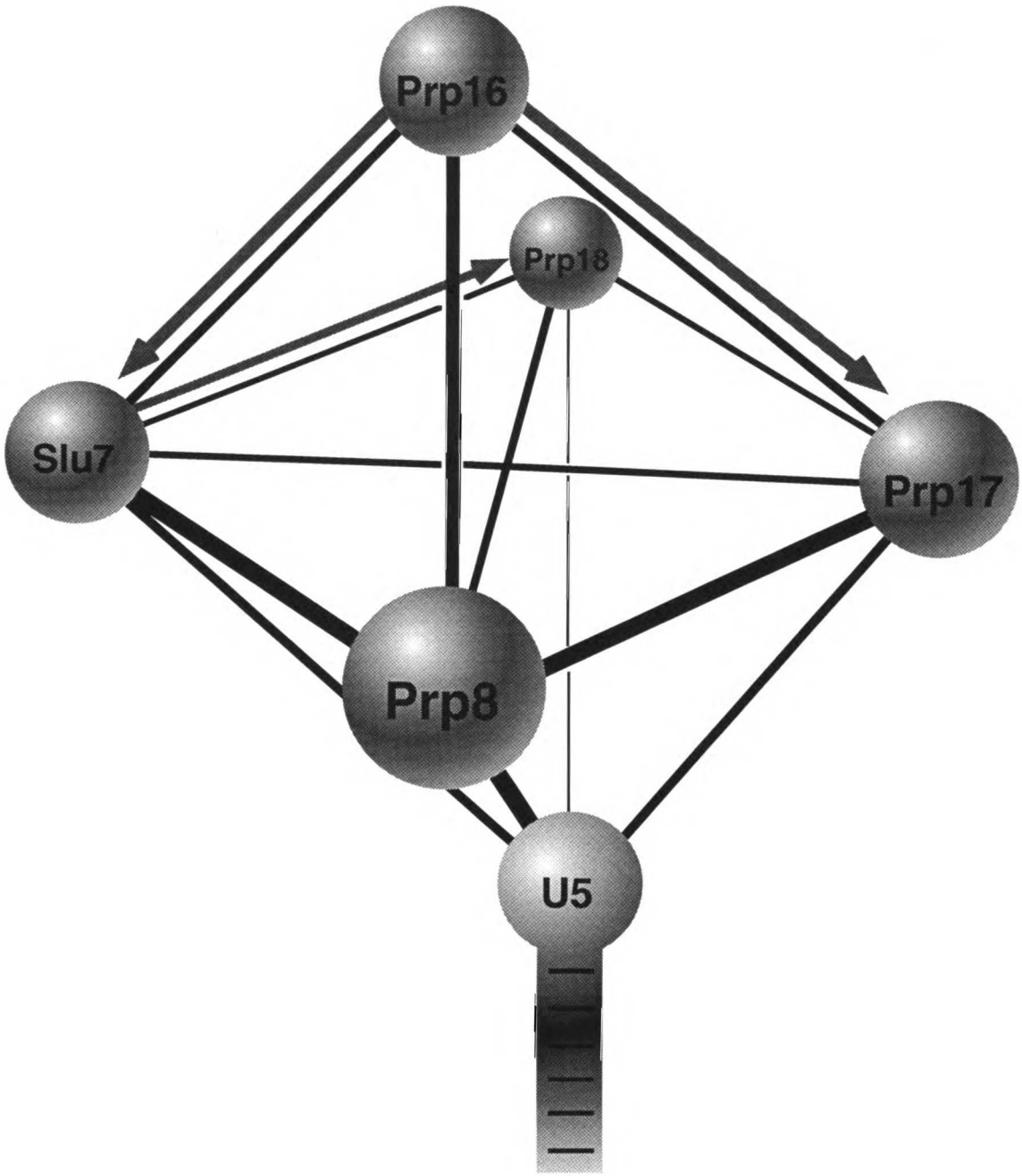
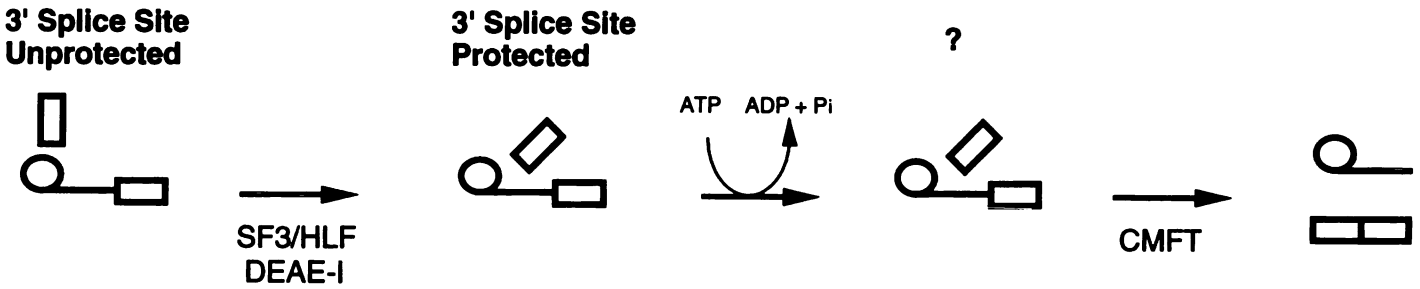

Prp16
Prp18
Slu7
Prp17
Prp8
U5

Figure 6. Conformational rearrangements at the 3' splice site during the second catalytic step. The lariat intermediate is cartooned with the 3' splice site in an unprotected (exons perpendicular) or protected (exons at acute angle) state. Protein factor and ATP requirements are listed under the arrows that show the transitions between stages. A question mark indicates that no information is available. 3' splice site crosslinking is shown as spheres for yeast proteins Prp16 (16), Slu7 (7), and Prp8 (8). The equilibrium depicted in the unprotected state for yeast represents strong 3' splice site crosslinking of Prp16 and weaker crosslinking of Prp8 and Slu7. The protein factors are described in Table 1.

Mammals



Yeast

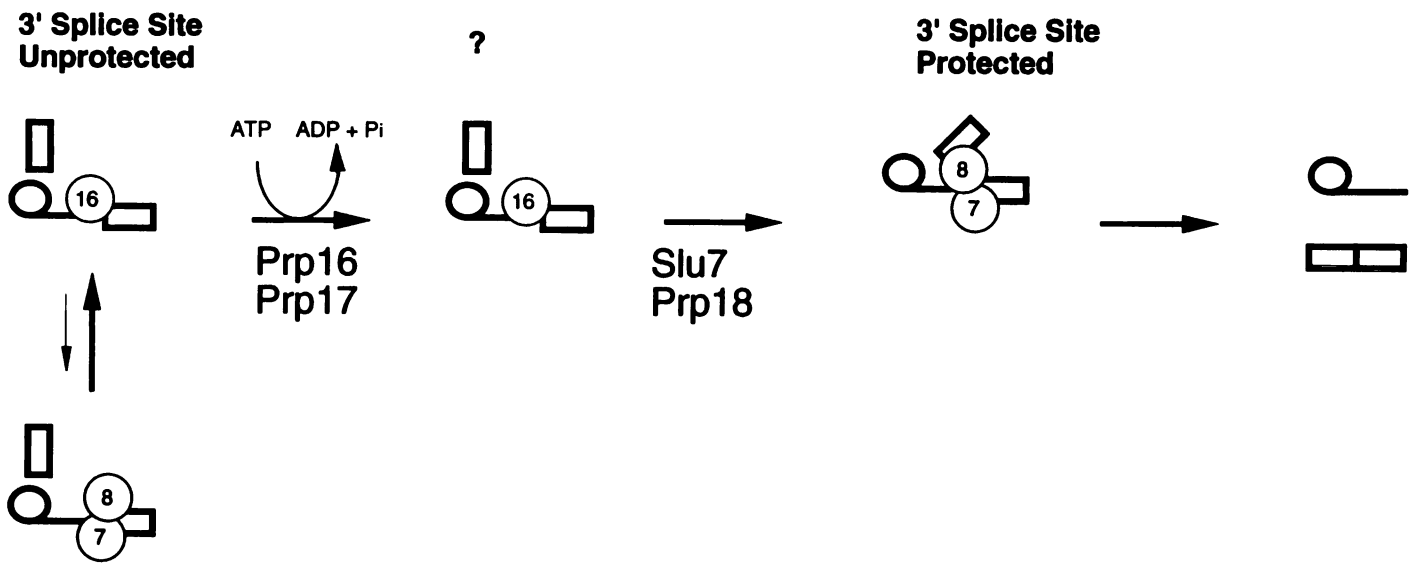


Figure 7. Spliceosomal versus group II putative catalytic domains. Top. Originally proposed analogy between yeast U2/U6 helix I (left) and consensus nucleotides for group IIA domain V (right). Step two specific residues are in uppercase bold. The conserved AGC trinucleotide is shaded in each structure. Non-bridging phosphate oxygens required for the first (filled) and second (open) steps are indicated by the lollipop shapes. Bottom. Revised analogy which extends to include an intramolecular stem in U6. The data on U6 phosphates are combined from yeast and nematodes. See text for citations.

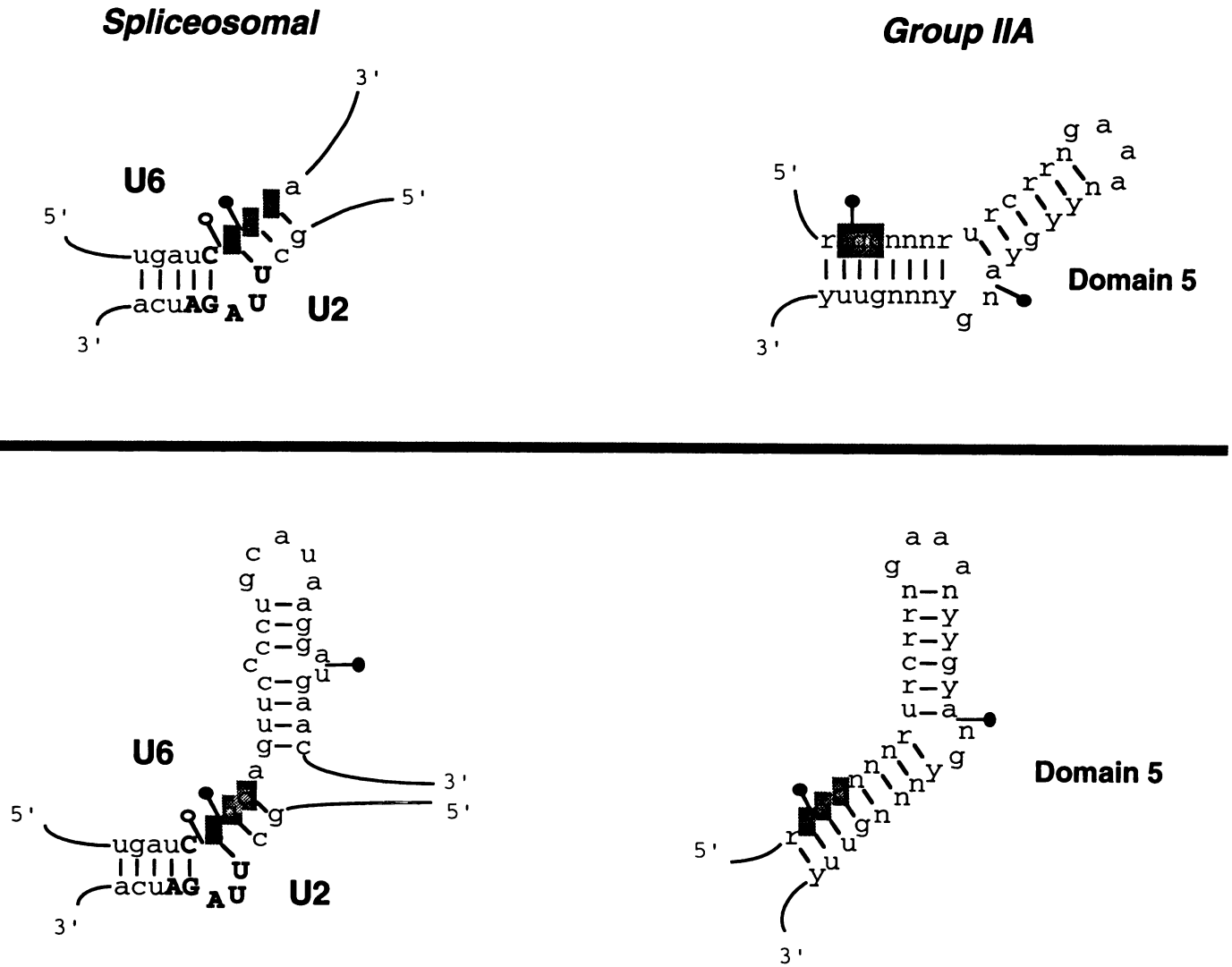

Spliceosomal
U6
U2
Group IIA
Domain 5
Domain 5

Table 1.

a. Indicates whether the encoded gene product is essential for viability in yeast. na, not applicable. nd, not determined

b. Indicates the order of protein function with respect to the ATP requirement during step two. Yes indicates that the protein functions at or before an ATP-requiring step. No indicates that the protein functions in the absence of ATP. nd, not determined.

c. These proteins are complex mixtures which have been given no designation by the authors. The column fractions from which they are derived are used as a provisional means of nomenclature.

Table 1.
Proteins Required for the Second Step of Splicing.

Protein	MW (kD)	Essential?[a]	Sequence Motifs	ATP Requirement?[b]	Functions	References
Yeast Proteins						
Prp16	120	Yes	DEAH	Yes	RNA-dependent ATPase; Branchsite suppressor; Induces conformational change at 3' splice site; Crosslinks to 3' splice site	Couto et al., 1987; Burgess et al., 1990; Schwer & Guthrie, 1991; Schmid & Linder, 1992; Schwer & Guthrie, 1992; Burgess & Guthrie, 1993; Umen & Guthrie, 1995c
Prp17	52	No	WD	Yes		Neer et al., 1994; Jones et al., 1995
Prp18	28	No		No	U5 snRNP associated	Horowitz & Abelson, 1993b; Horowitz & Abelson, 1993a
Slu7	44	Yes	Zinc Knuckle	No	3' splice site selection; Crosslinks to 3' Splice site	Frank & Guthrie, 1992; Ansari & Schwer, 1995; Jones et al., 1995; Umen & Guthrie, 1995c
Prp8	280	Yes	Proline-rich domain	nd	U5 snRNP protein; 3' splice site selection; Crosslinks to 5' splice site, branch site and 3' splice site; Required for first step; Required for U4/U5/U6 snRNP stability	Lossky et al., 1987; Jackson et al., 1988; Brown & Beggs, 1992; Hodges et al., 1995; Teigelkamp et al., 1995; Umen & Guthrie, 1995b; Umen & Guthrie, 1995c; Umen & Guthrie, 1995a
Ssf1	~200	nd		nd	Functions after Prp16	Ansari & Schwer, 1995
Mammalian Proteins						
PSF	100	na	RGG; RRM; proline/ glutamine-rich	nd	Pyrimidine tract binding	Patton et al., 1993; Gozani et al., 1994
SF3/HLF	nd	na		Yes	Heat labile	Krainer & Maniatis, 1985; Sawa & Shimura, 1991
SF4a	nd	na		nd		Krainer & Maniatis, 1985
CMFT[c]	nd	na		No		Lindsey et al., 1995
DEAE-I[c]	nd	na		Yes		Sawa & Shimura, 1991
Ia	nd	na		nd		Perkins et al., 1986

APPENDIX 1

RNA-Protein Crosslinking With Site-Specifically Labeled Substrates

INTRODUCTION

With the development of a technique for high efficiency sequence-specific RNA ligations, site specific crosslinking is now possible with virtually any RNA substrate. Sequence specific RNA ligation allows construction of RNAs of any virtually any size with single modified and/or labeled nucleotides at defined internal positions (Moore & Sharp, 1992). This technology is ideal for examining RNA-RNA and RNA-protein interactions in complex systems such as the spliceosome or ribosome. The advantages of this method over conventional binding assays are several fold. First, the interaction of interest can be examined in its native context i.e. in a splicing reaction. Second, with the use of single phosphate labels, interactions can be automatically mapped with nucleotide resolution. Third, transient or low affinity interactions that might not be detectable with a conventional binding assay can be captured with site-specific crosslinking. The technique has already been applied, with great success, to identify protein-RNA and RNA-RNA interactions in the spliceosome (Wyatt et al., 1992; Sontheimer & Steitz, 1993; MacMillan et al., 1994; Sontheimer, 1994; Teigelkamp et al., 1995; Umen & Guthrie, 1995a; Umen & Guthrie, 1995b). For other useful reviews see (Hanna, 1989; Sontheimer, 1994).

While this methodology is very powerful, it is important to bear in mind its limitations. While the detection of a crosslink is evidence of a specific binding interaction between two molecules, crosslinking is not the same as binding. In theory, two molecules can be close enough in space to allow a crosslink without any binding energy being associated with their proximity. This becomes an increasing problem as the radius of crosslinking increases. The radius of crosslinking is the farthest distance from the nucleotide of interest where one can obtain a crosslink. This parameter is influenced by several factors. First is the location of the photoactivatable group with respect to the nucleotide. In many cases, the nucleotide itself is photoactivatable and is considered a zero length crosslinking group e.g. 4-thio-uridine. In other cases the photoactivatable

nucleotide is on an arm of variable length which increases the radius considerably (MacMillan et al., 1994). It also decreases the probability that a crosslink represents a direct interaction between the nucleotide and the crosslinked adduct. A second parameter that influences the radius of crosslinking is the half life of the photoactivated species. The longer the half life, the greater the chance of a random diffusion event generating a crosslink. For example, a protein bound to a nearby segment of RNA might have a high enough local concentration to collide with the activated nucleotide often enough to see a crosslink. A third, and related, parameter involves the overall rigidity of the complex one is examining. In a very flexible or dynamic structure, the environment of the activated nucleotide can change rapidly, exposing it to multiple transient interactions. Some of these interactions may be non-specific and obscure the relevant ones or complicate analysis. The latter two parameters are difficult or impossible to measure and thus, the true radius of crosslinking is also difficult to accurately determine for a given type of experiment. Therefore, it is important to utilize independent methods to critically evaluate the significance of a particular crosslink.

SUBSTRATE DESIGN

Theoretically, site-specific RNA ligations allow a labeled photoactivatable nucleotide to be placed anywhere in an RNA. On a practical level, there are some limitations based on available reagents and the specificities of the enzymes involved. To date, uridine is the only nucleotide with commercially available photoactivatable forms that are incorporated by bacterial RNA polymerases and are not blocked at their 5' or 3' ends. The most popular forms are 5-bromo-UTP (Sigma), 4-thio-UTP (USB/Amersham) and 4-thio-UpG (Sigma). The latter two are useful for identifying both RNA-RNA and protein-RNA crosslinks. Note that the first two analogs do not automatically get incorporated site specifically. This is addressed below. Alternatively, specific analogs are not absolutely necessary for crosslinking since nucleotides themselves are

photoactivatable with 254 nm light. However, analogs improve the efficiency of photoactivation considerably. Furthermore, the higher wavelengths of UV light used to excite the analogs is less damaging to proteins.

Before a substrate can be designed, the target sequence of interest must be identified. Although single nucleotide resolution is possible, it is not always desirable. For example, if one were looking for a protein-RNA interaction at the intron branchsite, certain crosslinks might be missed if only a single nucleotide were photoactivatable. An alternative is to design a labeled fragment with multiple photoactivatable groups. RNase T1 is a useful enzyme for this purpose since it cuts after G residues leaving a 3' terminal phosphate group. Thus a small labeled RNA oligo that is flanked by two G residues (e.g. Gp/NNNNNNNNGp/) and contains one or more modified uridine residues could be the target RNA. The ligation substrate should be designed in two parts such that the 3' terminal region of the 5' half RNA ends with the RNase T1 oligo sequence of interest. The 3' half molecule will usually start with a G residue to accommodate T7 RNA polymerase (T7 RNAP) and then continue with the remainder of the RNA sequence. For the oligo of interest to become labeled, the downstream (3' half) RNA must first be labeled with polynucleotide kinase so that the 5' end contains 32P. After ligation and RNase T1 digestion, this label will be transferred to the last G residue of the upstream RNA oligo. With this method, the entire RNA molecule can contain photoactivatable nucleotides. The specificity is introduced from the 32P label which is only on a single RNase T1 fragment in the digested RNA. Any protein that becomes labeled after crosslinking will have to have come in contact with that short RNA fragment. A smaller RNA oligo or single nucleotide can become the target simply by using an RNase with broader specificity such as RNase P or RNase A. One must pay attention to which side of the phosphodiester bond the RNase cleaves in order to ensure that the labeled phosphate is associated with the correct oligo or nucleotide after digestion. Alternatively, a three-way RNA ligation can be used. The middle RNA is an oligomer which is

synthesized chemically or enzymatically. Only the oligomer need contain the modified nucleotide and P32 label.

PREPARATION OF RNAS

1. Template design

Site specific RNA ligations require large amounts of RNA (1-2 μm in a 10-20 μl reaction) with homogenous ends. Therefore, care must be taken in the design and synthesis of the ligation substrates.

First, the T7 RNAP or SP6 RNAP promoter sequence should be as close to optimal as possible. T7 RNAP has a sequence preference of GGGAG for the first five templated nucleotides (Milligan & Uhlenbeck, 1989; Yisraeli & Melton, 1989). The first G is critical for reasonable yields and the second is important. Increased amounts of T7 RNAP in the transcription reaction can help overcome a non-optimal 5' end.

Large amounts of transcription template must be prepared. While this can be accomplished by cloning the RNA of interest into a conventional vector such as Bluescript, PCR is a much better method for template preparation. A template sequence of 300 nucleotides in a 3 kb vector makes up only ~10% of the total DNA mass. One μg of this template is required where 100 ng of the same template made by PCR could be used. This is important because DNA tends to precipitate in T7 transcription buffer at moderate concentrations (.4-.5 μgms/μl). A more important reason for using PCR to generate the transcription template is that the ends of the RNAs can be precisely defined. This is a prerequisite for sequence specific RNA ligations.

For the PCR method, a primer must be designed with a T7 promoter in the upstream (5') oligo for each RNA. Thus, the 5' PCR primer should contain the sequence (NNNNTAATACGACTCACTATAgggagN'N'N'N' . . . .) where optional nucleotides are denoted N, the promoter sequences are in uppercase (non transcribed) or lowercase

(transcribed; partly optional [see above]) and the homology with the target DNA begins with these lower case nucleotides and continues with the N' sequences.

In principle, the downstream (3') oligo for each RNA can simply match the size and sequence of the desired 3' end. If this is done, the PCR reaction should be treated with T4 DNA polymerase or the Klenow fragment to make the template ends flush. A better method, if possible, is to use a restriction enzyme to generate the 3' end. Only enzymes that create blunt ends or 5' overhangs should be used because 3' overhangs can result in significant snap-back synthesis by T7 RNAP. The advantage of using restriction digestion is that the 3' end will be absolutely uniform. The only problem might be that the 3' end sequence that is desired cannot be generated with known restriction enzymes. However, the multiplicity of restriction enzyme recognition sequences makes a match highly probable. If a restriction enzyme is used to create the 3' end, a sufficient number of extra bases (4-6) should be added to the downstream oligo after the restriction site so that digestion is not inhibited by DNA end effects.

2. PCR protocol for template generation

To generate microgram quantities of short PCR templates, large scale PCR reactions (400-500 µl) are utilized. The reaction is divided into smaller aliquots (50-100 µl) to accommodate the volume and tube size limitations of most thermocyclers. A typical reaction is as follows:

40 µl primer 1 (10 µM)
40 µl primer 2 (10 µM)
40 µl 10X Hot Tub DNA polymerase buffer
32 µl dNTP mix (2.5 mM each dNTP)
40 µl target DNA (1ng/µl)
4 µl Hot Tub DNA polymerase (12 units)
204 µl H2O
400 µl

Amplify for 25-30 cycles and remember to use a lower annealing temperature for the first cycle since the primer with the T7 promoter will not fully anneal in the first

round. After the PCR reaction extract the DNA twice with phenol/chloroform and precipitate with .4 volumes 3M NaOAc (160 µl) and 1.4 volumes of isopropanol (560 µl). Digest DNA to completion with restriction enzyme of choice, keeping in mind that the concentration of restriction sites per unit mass of DNA will be very high. After restriction digestion, extract and precipitate the DNA as above. Isopropanol does not allow nucleotides and small oligos to precipitate efficiently so the DNA does not need to undergo any further purification. At this point, the concentration can be measured by UV absorption. Expect approximately 12-16 µgms of a 100 base pair DNA per 500 µl PCR reaction.

3. Transcription reactions

High yields of full length RNAs can be prepared using a kit produced by Ambion (T7-MEGAshortscript) or with easily obtainable reagents (Milligan & Uhlenbeck, 1989; Yisraeli & Melton, 1989). The kit is economical to use once the conditions for making and ligating the RNAs of interest have been determined. NTP stocks should kept at -80° for long term storage but small aliquots can be stored for several months at -20°. A typical reaction contains the following:

5-10 µCi 3H-UTP dried down (used as a tracer to calculate yield)
2 mM each ATP, GTP and CTP
2 mM total (UTP + modified UTP)
1x T7 transcription buffer (40 mM Tris pH 8, 6 mM MgCl2, 1 mM spermidine, .01% Triton X-100)
8 mM additional MgCl2 (to match NTP concentration)
5 mM DTT (make from concentrated (1M) stock each time)
200-400 nM PCR template (usually 1-3 µgms)
1 µl Promega recombinant RNasin (optional)
2.5 units inorganic pyrophosphatase (Sigma; 2.5 U/µl in 50% glycerol, 20 mM KPO4 pH 7, .1 mM DTT, .1mM EDTA; store at -20°)
800-1000 units of T7 RNAP (USB or Pharmacia)
DEPC H2O to 100 µl.

Incubate at 37° for 3-6 hours. An incubator is better than a water bath since the mixture does not evaporate.

Notes: The 3H-UTP is a good tracer for calculating yields since it can't be visualized by standard autoradiography. For BrUTP, 50-100% can be incorporated as modified nucleotide. 4ThioUTP inhibits the transcription reaction and must be titrated. If 4thio UpG is used, it must be present in excess over GTP to ensure that the majority of transcripts initiate with the modified dinucleotide. Cap (GpppG) is not absolutely necessary for the stability of RNAs in yeast splicing extracts and is omitted. However, capped RNAs have been reported to splice better than uncapped ones. Pyrophosphatase increases yield by eliminating one of the transcription products.

After transcription, add 1 µl of Promega RQ1 RNase-free DNase and incubate 20-30 minutes more.

For the 5' half RNA, add EDTA to 20 mM and extract the reaction with phenol/chloroform. Precipitate with .1 volume 3M NaOAC and 2.5 volumes ethanol. For the 3' half RNA, the phosphate must be removed from its 5' end. Pass the transcription reaction over two 1 ml G50 spin columns equilibrated with phosphatase buffer (50 mM Tris pH 8.4, .1 mM EDTA). Add 1 µl (~20 units) alkaline phosphatase (Boehringer, high concentration, molecular biology grade) and incubate 30 minutes at room temperature. Heat to 95° for 2 minutes and extract twice with phenol/chloroform. Precipitate as above. Organic extraction is critical for allowing resuspension of the RNAs after precipitation. Remember to keep all tubes and supernatants for special disposal as 3H waste.

Purify the RNAs in a denaturing polyacrylamide gel keeping in mind that a lane on a 20 ml, 20 cm X 20 cm gel (15 well comb) can hold approximately 10-20 µgms RNA before it is overloaded. Visualize the RNA by UV shadowing (be brief so as not to photoactivate modified nucleotides). To UV shadow, place the gel on a piece of Saran Wrap and set on an intensifying screen or fluorescent TLC plate (The glass gel plate will block the UV light so it must be removed). Turn off lights and visualize by shining a 254 nm UV light on the gel. Cut out and elute the RNA (dark band) from the crushed gel

slice by nutating 4 hours-overnight at 4° with .3M NaOAC in TE with an equal volume of phenol/chloroform. Precipitate as above and quantitate yield. 20-50% incorporation of UTP is typical for final yields. Store the RNAs in aliquots at -80°.

4. Labeling the 3' half RNA

The 3' half RNA must be labeled at its 5' end prior to ligation. This is done with an aliquot of the RNA just prior to ligation. Prepare the following reaction:

3' half RNA (enough for a 10 µl ligation reaction at 2 µM)
4 µl 10X polynucleotide kinase buffer (USB)
3 µl fresh gamma 32P-ATP (480 µCi)
.5 µl (5 units) T4 polynucleotide kinase (USB)
H2O to 40 µl

Incubate at 37° for 20'. Add 1 µl 20 mM cold ATP and incubate 10 more minutes at 37°. The cold ATP completes the reaction and prevents dead end ternary complexes from forming in the subsequent ligation. Increase volume to 100 µl with TE. Separate RNA by passing over two 1 ml G50 spin columns equilibrated in TE. Extract once with phenol/chloroform and ethanol precipitate. Prior to precipitation, count 1 µl to determine kinasing efficiency . Typical yields are 5 X 10^7 to 10^8 cpm.

5. Ligations

Ligation conditions have been described (Moore & Sharp, 1992; Wyatt et al., 1992; Sontheimer, 1994). The kinased 3' half RNA is centrifuged, washed with 70% ethanol and dried very briefly (1-2 minutes) in the speed vac. The RNA is resuspended with 5' half RNA and bridging oligonucleotide (also gel purified) so that the final concentrations in a 10 µl reaction will be 2 µM RNAs and 1.5 µM oligo. The bridging oligo must anneal stably to each half. A 40-50° Tm for each half of the substrate is reasonable for this purpose. The oligo should never be added in excess of the RNAs since it will then form binary complexes rather than ternary complexes, and thus act as an inhibitor of the ligation. The mixture is heated to 95° for two minutes and left at room temperature for 5-10 minutes. H2O, 1 µl 10X ligation buffer (.5M Tris pH 7.5, 10 mM ATP, 10 mM

MgCl2, 1mg/ml BSA, .4M DTT), 1 µl USB high concentration ligase and .5 µl Promega recombinant RNasin (10 µl final volume) are added and the reaction is incubated at 37° for 1-2 hours. Only USB ligase should be used since they purify away ribonuclease activity. After 1 hour, the reaction can be reheated and RNasin and ligase readded. A second aliquot of DNA ligase can be added whether or not the reaction is reheated and cooled. Note that contaminating RNases can easily ruin the reaction at this step. RNase usually copurifies with RNasin and can be released if the RNasin is old or is heated. The advantage of the heating and recooling step is that the ternary complexes of RNAs and oligo can reassort so that dead end complexes are disassembled and active ternary complexes can reform. The reheating step must be tested empirically with each ligation substrate to determine whether it increases or decreases the yield of ligated RNA.

Gel purify the ligated RNA by adding an equal volume of formamide loading buffer and loading the sample directly on a denaturing polyacrylamide gel (as above). When the ligated RNA has almost migrated to the desired location, 1% of the kinased 3' half RNA can be loaded into a

different lane and run into the gel several centimeters. This can be used as a visual standard to estimate ligation efficiency. i.e. if the ligation product is ten times brighter than the 1% standard, the ligation efficiency is 10%. Ligation efficiencies of 10-50% are typical and depend on conditions and the nature of the substrates.

CROSSLINKING REACTIONS

The general protocol for site specific crosslinking is to incubate the substrate in extract, treat with UV light, RNase digest and immunoprecipitate. The crosslinked protein(s) is then fractionated by SDS-PAGE, transferred to nitrocellulose and then visualized by autoradiography and immunoblotting.

For splicing reactions, the amount of substrate should be titrated. If too little is added, the crosslink might not be visualized. If too much is added, the reaction can be

inhibited. More than 10% of a ligation reaction in a 40 µl splicing reaction can be inhibitory. This puts an upper limit on the amount of substrate that can be added. For the Prp8p-, Prp16p- and Slu7p-3' splice site crosslinks, 3 X 10^5 cpm substrate gives a reasonable signal.

After incubation, 5-10% of the reaction should be set aside for RNA analysis. The remainder is transferred to a U-well microtiter dish (Falcon, Cat. # 3911) that has been pushed into a bed of crushed ice. The correct type of microtiter dish is important since some types of plastic will themselves efficiently crosslink to RNAs. A hand held UV lamp is placed directly on the microtiter dish over the sample which is irradiated for 1-30 minutes. The location of the bulb in the lamp must be determined ahead of time so that the bulb is directly over one row of wells in the microtiter dish. All samples should be placed in that row. For nucleotides alone, use 254 nm light. For 5-bromo-uridine, use 312 nm light and for 4-thio-uridine use 365 nm light. The time of irradiation must also be titrated. As time increases, the amount of crosslinking increases and then plateaus. At still later times, crosslinking appears to decrease. The decrease is due to the fact that proteins in the reaction become photodamaged and denature. This is a much greater problem with 254 nm light than with 312 nm or 365 nm light. Nonetheless, it is prudent to irradiate the samples as briefly as possible.

After crosslinking, the samples are transferred back into an eppendorf tube and protease inhibitors (1 mM PMSF, 1 µg/ml leupeptin, 1 µg/ml aprotinin, 1 mM benzamidine, 5 mM EDTA) and RNase are added. The microtiter dish can be monitored to make sure that most of the RNA is not stuck to the dish. The samples are then incubated at 35° for 30 minutes to allow digestion of the RNA. The incubation temperature should not be raised above 37° since significant protein denaturation begins to take place at these temperatures. The RNase should also be titrated so that the minimal amount that gives complete digestion is used. With too much RNase, cleavages take place at nonspecific nucleotides (i.e. A, C and U with RNase T1). 10 units/µl is a good

starting concentration for RNase T1 (Boehringer). At this point, a portion of the samples can be removed and mixed with protein sample buffer to examine total protein crosslinking.

After digestion, the reactions are centrifuged at 4° for 5 minutes to remove precipitates. The supernatants are then incubated with antibodies on ice or added to protein A sepharose beads prebound with antibodies in NET 150 buffer (150 mM NaCl, 50 mM Tris 7.4, .05% NP40). For the latter procedure, protease inhibitors should be added to the mixture. After antibody binding, the reaction is centrifuged again and the mixture is transferred to protein A sepharose beads in NET150 as above. After, the complexes are bound to the protein A sepharose, the beads should be washed carefully and thoroughly with NET 150 buffer 3-4 times in a volume of .5-1 ml. Concentrated (3X-4X) protein sample buffer is added to the pellets so that the final supernatant volume is minimal (20-30 µl). The samples are boiled immediately for two minutes and stored at -20°. They are reboiled for 5 minutes prior to electrophoresis.

After SDS-PAGE, the gel can be immediately dried and visualized but it is better to transfer the proteins to nitrocellulose using a standard western transfer procedure. In this way, after autoradiography, the nitrocellulose filter can be immunoblotted and the comigration of the crosslinked protein and the immunoreactive protein can be confirmed. Furthermore, if the crosslinking result is negative, the success of the immunoprecipitation can be evaluated. All autoradiography (usually phosphorimage analysis) should be completed before immunoblotting since nucleases in sera can remove crosslinked RNA fragments.

CONTROLS AND INTERPRETATION

Once a crosslinking signal is obtained for a protein, its relevance must be evaluated. Controls are performed to ensure several minimal criteria.

First, the crosslinked protein should be positively identified. This can be accomplished through immunoblotting as described above. If the protein is epitope tagged, an untagged version should be used. Alternatively, a mock reaction with pre-immune serum or no antibodies can be performed in parallel to demonstrate the specificity of the immunoprecipitation.

Second, the reaction should be performed without UV irradiation and/or without the photoactivatable nucleotide incorporated to demonstrate that a covalent crosslink is being observed. As mentioned above, RNase digestion should also be confirmed, either with the entire sample or, ideally, with the crosslinked immunoprecipitated protein. The pellet from the immunoprecipitation is protease treated and extracted. The liberated RNA oligonucleotides are then precipitated and examined on a 15-20% denaturing gel. The RNA will migrate slightly slower than expected since there are attached peptide fragments.

The specificity of the crosslink should be examined. Without ATP, many splicing-specific crosslinks will disappear. The same will be true with some inactivated *prp* mutant extracts or with substrate mutations that block splicing. The site specificity of the crosslink can be shown by demonstrating that no crosslink is observed when the label is placed elsewhere in the substrate.

Finally, if a protein-RNA crosslink meets the above criteria, it must be placed in the context of other data to be truly informative. Either genetic or biochemical experiments that suggest a functional interaction with the RNA of interest can provide strong confirmatory data supporting the relevance of the crosslink.

APPENDIX 2

Identification of a U1 Suppressor Mutant That Encodes a Putative Splicing Factor With Homology to Ubiquitin Conjugating Proteins

ABSTRACT

Yeast U1 snRNA is several fold larger than its metazoan counterparts. The extra length of yeast U1 is accounted for by a large yeast-specific domain inserted into a secondary structure similar to that of metazoan U1s. Although deletion of most of the yeast-specific domain is tolerated, cells containing deletions in this region display slow growth and temperature sensitive phenotypes. A temperature resistant suppressor of one such deletion mutant was isolated. The suppression phenotype results from two independent mutations in genes *SSR1* and *SSR2* (Suppressor of snRNA). The combination of *ssr2-1* and *ssr1-1* can suppress several types of deletions in U1, including one deletion that is lethal in a wild type background. One of these mutants, *ssr2-1*, appears to independently confer a splicing phenotype in the presence of a wild type U1 allele, and itself causes temperature sensitive growth. *ssr2-1* was found to be allelic to *slu12-1* which was isolated as a synthetic lethal with a U5 snRNA loop mutation. *ssr2-1* and *slu12-1* also display unlinked non-complementation with *prp3-1*, a U4/U6 snRNP protein. Whereas extracts prepared from U1 deletion strains block spliceosome assembly at an early stage prior to "committment", *ssr2-1* extracts display a block at a later stage and appear to assemble complete spliceosomes. A putative clone of *SSR2* and two high copy suppressor plasmids have been isolated. The sequence of the complementing region from *SSR2* reveals it to encode an unusually large protein (3268 residues) with homology in its C-terminal domain to a family of ubiquitinating enzymes.

RESULTS

A deletion mutant of the U1 encoding gene *SNR19* termed Δ10 removes 196 nucleotides of U1 sequence in the yeast-specific domain (Kretzner et al., 1990)(nts. 248-444) and contains a 16 nt insertion of two BamHI linkers (Paul Siliciano, unpublished; Figure 1). This allele causes slow growth at 25°-34° and inviablity at 37° (Figure 2). Previous work has demonstrated that deletions in this region of U1 do not significantly affect the steady state levels U1 RNA (Siliciano et al., 1991).

In order to identify splicing factors that interact with U1, we isolated suppressors of the temperature sensitive growth phenotype of a U1 Δ10 strain. This strain is deleted for its chromosomal copy of U1 and contains the Δ10 allele on a plasmid. The Δ10 strain was grown in YEPD and mutagenized with UV radiation to kill 80-90% of the cells. ~10^8 viable mutagenized cells were spread on 10 YEPD plates and incubated at 37° for 10 days in the absence of light (to prevent photorepair). A single colony (Sup1) was isolated and purified by restreaking the cells at 37°.

The Sup1 strain was mated to an isogenic parent strain containing U1 Δ10 and the resultant diploid was sporulated. The growth phenotype of the Sup1/+, U1 Δ10 strain resembled that of the Δ10 strain although very slight growth was observed at 37°. Therefore, Sup1 behaves as a recessive suppressor. Spore progeny from the Sup1/+ cross were dissected and analyzed for growth at 30°, 34° and 37°. Instead of the expected segregation pattern of 2 suppressor: 2 wild type that should be observed if the Sup1 strain contains a single suppressor locus, a more complicated pattern was observed. Approximately 1/4 of the spore progeny behaved like the original Sup1 strain and were viable at 37°. Approximately 1/2 of the spore progeny behaved similar to the parent strain (U1 Δ10) and were viable at 34° but dead at 37°. Approximately 1/4 of the spore progeny were sick but viable at 30° and dead at 34° and 37°.

This segregation pattern suggested that suppression was due to the action of two independent mutations. We hypothesized the existence of two mutant alleles, *ssr1-1* and

ssr2-1 that are both required for suppression. In addition, we surmised that the *ssr2-1* allele, in the absence of *ssr1-1*, itself conferred a temperature sensitive phenotype. Somewhat paradoxically, this mutation exacerbates the phenotype of U1 Δ10 strains in the absence of *ssr1-1* (see Table 1). To test the double mutation hypothesis, we assigned putative genotypes to spores from a tetratype and parental ditype tetrad and recrossed these to the parent strain. The phenotypes for each genotype are predicted to be as follows: *ssr1-1 ssr2-1*, 34° + 37° +; *ssr1-1 SSR2*, 34°+ 37°-; *SSR1 ssr2-1*, 34°- 37°-; *SSR1 SSR2*, 34°+ 37°-. As predicted, the putative *ssr1-1 ssr2-1* spore gave the same complex segregation pattern as the original Sup1 strain upon backcrossing (data not shown). The *ssr1-1 SSR2*, and *SSR1 SSR2* strains could not be distinguished phenotypically, and both yielded spores with only parental phenotypes (U1 Δ10-like) when backcrossed (data not shown). Importantly, the *SSR1 ssr2-1* strain displayed a simple 2:2 (34° + 37° -: 34°- 37°-) segregation pattern when backcrossed to the U1 Δ10 parent (data not shown). Finally, a *ssr1-1 SSR2* strain was identified from a non-parental ditype segregant and crossed to a *SSR1 ssr2-1* strain. As expected, the spore progeny displayed the same segregation pattern as the original Sup1 strain (data not shown). Thus, the results from backcrossing fulfilled the predictions of a two locus model for suppression. The growth phenotypes of these strains are shown in Figure 2 and Table 1.

To test the allele specificity of the U1 suppressor mutations, we replaced the Δ10 allele with either wild type U1 or various other U1 alleles (Siliciano et al., 1991)(Table 2). For most deletion mutants (described in table 2), we found that the combination of *ssr1-1* and *ssr2-1* can suppress their temperature sensitive phenotypes. In addition, either *ssr2-1* alone or in combination with *ssr1-1* can suppress a lethal deletion mutant, Δ27, allowing growth at 18° but not at 30° or higher (Table 2). Thus, *ssr2-1* and *ssr1-1* can suppress a range of deletions in the non-conserved domain of yeast U1. In contrast, we found that the phenotypes of point mutations in the 5' end of U1 are excacerbated by both

ssr1-1 and *ssr2-1*. Thus, these mutations interact both positively and negatively with different U1 alleles.

Since *ssr2-1* independently confers a temperature sensitive growth phenotype, we were able to test for complementation with known splicing mutants. We found that *ssr2-1* does not complement a novel, unpublished mutant, *slu12-1* (Figure 3; see Dan Frank's dissertation; synthetically lethal with U5 C97,99). To determine whether these mutations are allelic, a *ssr2-1/slu12-1* diploid was constructed and sporulated. Of 8 four spore tetrads analyzed, all the spore progeny were temperature sensitive indicating that *ssr2-1* and *slu12-1* are linked and almost certainly allelic (data not shown). This is supported by the finding that both these mutants display unlinked non-complementation with *prp3-1* at 37°. When *prp3-1 +/+ ssr2-1* or *prp3-1 +/+ slu12-1* diploids were constructed, they exhibited poor growth at 37° on synthetic media (Figure 3). Both segregation analysis and the fact that a *PRP3* plasmid does not complement the *ssr2-1* and *slu12-1* phenotypes indicates that these mutants are not alleles of *PRP3* (data not shown).

In order to better characterize the phenotypes of *ssr1-1* and *ssr2-1*, cells were grown at permissive (25°), or shifted to restrictive (37°) temperatures for 6 hours, and RNA was prepared and analyzed by primer extension. All combinations of *ssr1-1*, *ssr2-1*, and U1 Δ10 or wild type alleles were analyzed. In the presence of wild type U1 at restrictive temperature, the *ssr2-1* strain displays a loss of rp51a mature message, a characteristic of some splicing mutants (Figure 4, lane 6 versus 8). In the U1 Δ10 backround, all the strains show a strong loss of mature message phenotype with rp51a at the restrictive temperature (Figure 4, lanes 11-14 versus 15-18). Finally, in the presence of wild type U1, *ssr1-1* appears to suppress the splicing defect of *ssr2-1* (Figure 4, lanes 8 and 9). These experiments need to be repeated to ensure reproducibility and to examine other messages besides rp51a. Nonetheless, the data are consistent with *ssr2-1* causing impaired splicing at the restrictive temperature. Furthermore, the levels of U1 remain

constant in all lanes of Figure 4 indicating that the U1Δ10 defect and its suppression does not involve alterations in U1 RNA stability.

As a preliminary means of identifying the biochemical functions of Ssr1 and Ssr2, splicing extracts were prepared from each of the above strains that were used for primer extension analysis (8 total strains). In vitro splicing reactions containing labeled actin pre-mRNA were examined for each strain. The wild type strain, *ssr1-1* and *ssr1-1 ssr2-1* all produce moderate levels of splicing products (Figure 5, lanes 2, 3 and 5; most easily seen with the lariat intermediate) whereas *ssr2-1* splices poorly (lane 4). Since this strain displays a constitutive splicing defect, we did not attempt to heat-inactivate the extract. For strains containing U1 Δ10 alleles, splicing was uniformly poor and this defect is also constitutive. The strain containing *ssr2-1* in combination with U1 Δ10 showed the poorest splicing efficiency with no detectable intermediates or products (Figure 5, lanes 6, 7 and 9 versus lane 8).

In parallel, a portion of each of these reactions was examined on a complex gel to assay spliceosome assembly. As the gel system utilized has not been thoroughly characterized, the complexes are designated somewhat tentatively (Karen Shannon, personal communication). One problem that was encountered at this time was the appearance of slowly migrating spliceosome-like complexes in the absence of added ATP (Figure 6, lane 9). This has been seen in some extract preparations and may indicate the presence of residual ATP. Unfortunately, a reaction lacking ATP was an important control. The best evidence for the identity of the slowly migrating complexes as spliceosomes is that they are greatly reduced in abundance in the U1 Δ10 extracts. Furthermore, in other extracts, their presence has been observed to be ATP-dependent (Karen Shannon and Evi Strauss, personal communication).

For each extract, in the U1 Wt lanes, we see a fast migrating complex which is non-specific and three slowly migrating complexes which correspond to spliceosomes. The fastest of the three upper bands (B) is likely to contain U2 only, and the upper bands

(A1 and A2) are likely to contain fully assembeled spliceosomes based on analysis with other complex gel systems. The *ssr2-1* extract appears to be able to assemble all the splicing complexes similar to wild type, although kinetics have not been examined (Figure 6, compare lanes 1-4). Interestingly, in the U1 Δ10 extracts, the upper bands are largely missing and instead, there is a novel complex (C) that migrates above the non-specific complex. This is probably a blocked complex that is stalled at an early stage in the assembly reaction.

Since both *ssr2-1* and U1 Δ10 extracts display a strong block to splicing (Figure 5), we wished to determine whether they are blocked at the same or different steps. To accomplish this, the extracts were pre-incubated with labeled precursor for 10 minutes (Figure 7, First Addition). At this time, excess cold competitor RNA was added along with complementing extract (Second Addition). When the U1 Δ10 extract is pre-incubated as the first extract, no splicing can be observed upon complementation (Figure 7, lane 2). This indicates that the U1 Δ10 extract cannot form a committment complex. In contrast, the *ssr2-1* extract was fully complementable after pre-incubation followed by the addition of cold competitor RNA and U1 Δ10 extract (Figure 7, lane 3). The level of splicing observed in this experiment is similar to that seen when the U1 Δ10 and *ssr2-1* extracts are mixed at the beginning of the incubation period as a control (Figure 7, lane 1). Therefore, the defect in *ssr2-1* extracts is distinct from that seen in U1 Δ10 extracts. This result is also consistent with what we observed on complex gels i.e. U1 Δ10 extracts are blocked prior to spliceosome assembly.

In order to determine the nature of the *SSR2* gene product, we cloned the wild type locus by complementation of the temperature sensitive phenotype in a *ssr2-1* strain. Clones that rescued the growth defect at 34° or 37° were rescued by transformation into E. coli. These were then retested for suppression. In the process of screening multiple libraries for this low abundance clone, two high copy suppressors (Carlson and Botstein library) were also isolated. These suppressors do not fully complement the growth

phenotype of *ssr2-1* at 37° and have restriction maps that differ from the correct clone. When transferred to a low copy number vector, the level of complementation was further reduced. These clones have not been analyzed further. The putative correct clone displayed full complementation and was isolated from a low copy number library (Gift from M. Rose).

Partial restriction mapping of the correct clone indicated that part of the complementing activity resided within a 1.8 kb EcoRI fragment. However, the full length insert (~10-12 kb) could not be reduced in size without eliminating complementation. The 1.8 kb EcoRI fragment was sequenced and has been found to match a recently sequenced region of the yeast genome on chromosome IV (gp|U33050|SCD8035_1 D8035.1p). As expected from our inability to construct a minimal complementing clone, the presumptive *SSR2* open reading frame spans 9.8 kb and codes for a 3268 residue protein of predicted molecular weight 373 kd. There are two remaining expermiments that need to be done to characterize the *SSR2* clone. First, the complementing DNA must be integrated and shown to be linked to the *ssr2-1* mutation, and second, the gene must be disrupted to determine the null phenotype.

Using the 3268 amino acid sequence as a database probe, we found strong homology in the C-terminal region of Ssr2 to a large family of hect-domain-containing proteins (homologous to E6-AP carboxyl terminus), one of which is directly involved in ubiquitin transfer (E6AP)(Huibregtse et al., 1993; Huibregtse et al., 1995; Scheffner et al., 1995). Interestingly, the cysteine residue to which ubiquitin is conjugated as an intermediate in this process is conserved in Ssr2 and other family members (Figure 8). This suggests that the ubiquitin conjugation function has been conserved in Ssr2.

DISCUSSION

Using a temperature sensitive U1 allele (Δ10), we have isolated a suppressor strain that contains two mutations, *ssr1-1* and *ssr2-1*. Both of these mutations are

required for growth of U1 Δ10 at 37°. Quite surprisingly, *ssr2-1* alone exacerbates the temperature sensitivity of U1 Δ10 rather than suppressing it. *ssr1-1* appears to suppress the temperature sensitvity of *ssr2-1*, which in turn, allows suppression of the U1 deletion mutant. The suppression is not specific to U1 Δ10 but includes other of U1 deletions in the yeast specific domain. The fact that many of these alleles are suppressed suggests that the U1 deletion mutants all have a shared defect. Importantly, these alleles do not affect U1 snRNA stability and therefore, must affect U1 snRNP function (Siliciano et al., 1991; Figure 4). The defect could be in proper RNA folding and/or in snRNP protein association.

Two other genetic observations support the notion that Ssr2 is involved in splicing. First, an allele of *SSR2*, *slu12-1*, was independently isolated as a synthetic lethal with U5 snRNA. Second, both *ssr2-1* and *slu12-1* display unlinked non-complementation with a third splicing mutant, *prp3-1*. Prp3 is a U4/U6 snRNP-associated protein (Ruby & Abelson, 1991). Unlinked non-complementation is sometimes observed between alleles of genes that encode mulitsubunit complexes of proteins (Stearns & Botstein, 1988). Thus, Ssr2/Slu12 may be part of the U4/U6 or U4/U5/U6 triple snRNP.

The evidence for Ssr2 as a canonical splicing factor is, as yet, somewhat equivocal. In vivo analysis of splicing after a shift to the non-permissive temperature with an *ssr2-1* strain reveals a reduction in mature message. This phenotype has been seen with bona fide splicing factors such as *slu7-1* and *slu4-1* (Frank et al., 1992), but could also reflect a change in mRNA stability rather than a splicing defect. Analysis with more spliced and unspliced messages will be required to address this point.

In vitro, we find a constitutive splicing defect in *ssr2-1* extracts. Again, this precludes our ability to test for a direct role of Ssr2 in splicing by heat inactivation. It would be worthwhile to prepare extracts from *ssr2-1* and/or U1 Δ10 strains using a more reproducible procedure that has been developed recently (Umen & Guthrie, 1995). If

these extracts cannot be heat inactivated, then immunodepletion could be used to test the role of Ssr2 in splicing.

Interestingly, complex gel analysis suggests that spliceosomes can form in an *ssr2-1* extract, indicating a late role for Ssr2 in the splicing reaction. This observation is consistent with the previous observation of genetic interactions between *SSR2/SLU12* and U5 or *PRP3*. In contrast, U1 Δ10 extracts are blocked at an earlier stage. Once again, these experiments need to be repeated with a better characterized gel system. Nonetheless, ordering experiments support the complex gel analysis since the functional block in *ssr2-1* extracts can be distinguished from the block in U1 Δ10 extracts. *ssr2-1* extracts are blocked after committment of the message to the splicing pathway, whereas U1 Δ10 extracts are blocked prior to committment.

Finally, sequence analysis has yielded the surprising result that Ssr2 is very likely to be involved in a ubiquitination reaction. The homology to the E6AP family of ubiquitinating proteins is limited to the C-terminal ~350 amino acids (data not shown). The N-terminal regions of this protein family are completely divergent (Huibregtse et al., 1993; Huibregtse et al., 1995; Scheffner et al., 1995). It is likely that the N-terminal domains of this family are mediators of specificity for the ubiquitination reaction.

The results we have obtained seem somewhat paradoxical. How is it that a strong temperature sensitive mutation, U1 Δ10, is suppressed by another strong temperature sensitive mutation, *ssr2-1*? How is it that an early block in the splicing pathway is compensated for by a later block? And finally, how might ubiquitination fit into the splicing pathway? Without more evidence, we can only speculate on possible mechanisms of suppression. The model we propose here, if not correct in detail, may be an informative way to think about these results.

The fact that loss of function in one component of a pathway is compensated for by loss of function in other components is suggestive of antagonistic activities. Thus, Ssr2 may inhibit an activity of U1 snRNP. The two antagonistic activities must also be in

a rather delicate balance since a third mutation, *ssr1-1* is required to compensate for the loss of Ssr2 function. Ssr1 may therefore play a role in negatively regulating Ssr2. We further postulate that U1 Δ10 is blocked at two steps in the splicing pathway, an early step prior to committment, and a later step. Suppression may be occuring at the later step in which Ssr2 participates.

The fact that *ssr2-1/slu12-1* interact with both U1 and U4/U5/U6 suggests that the step in which Ssr2 participates is a link between the functions of these snRNPs. The most obvious link is at a hypothetical step in spliceosome formation where U5 and U6 replace U1 at the 5' splice site in a "hand-off" reaction (reviewed in Madhani & Guthrie, 1994; Nilsen, 1994). Mutant U1Δ10 snRNPs may be reluctant to let go at this point or get recognized improperly. In any case, loss of Ssr2 function would cause destabilization of the mutant U1 snRNPs, allowing U5 and U6 to bind the 5' splice site. This model also fits with the fact that *ssr2-1* exacerbates the effects of destabilizing point mutations in the 5' end of U1.

The putative biochemical activitiy of ubiquitination must also be fit into this model. The most obvious way to think about ubiquitination reactions is in protein degradation. However, we favor the scenario in which ubiquitin modifications are more stable. Stable ubiquitin modifications are not uncommon and could, in principle, be used like phosphorylation as a reversible post-translational modification to regulate protein activity (Hershko & Ciechanover, 1992). In which case, ubiquitination of either a U1 snRNP protein or a U4/U5/U6 snRNP protein would modulate 5' splice site binding. If the target were a U1 protein, ubiquitination would act to stabilize U1 binding to the 5' splice site. If the target were a U4/U5/U6 protein, ubiquitination would act to destabilize its binding. Given the genetic interaction between *prp3-1* and *ssr2-1*, Prp3 is a candidate for the target of Ssr2-dependent ubiquitination. Finally, Ssr1 might be a de-ubiquitination factor whose loss is required to compensate for lower ubiquitination rates in *ssr2-1* strains.

The above model makes testable predictions about the role of Ssr2 and ubiquitination in the splicing reaction. The most important prediction is that some splicing factors (e.g. Prp3) are ubiquitinated. Commercially available anti-ubiquitin antibodies (East Acres Biologicals) could, in principle, be used to test this prediction. The model also accomodates a direct role for Ssr2 in splicing without ruling out additional roles in other cellular processes. The connections established between Ssr2 and splicing thus far are strong enough to predict that whatever its mode of action, it will lead to interesting insights on the mechanism of splicing.

MATERIALS AND METHODS

Yeast Strains and Genetic Methods

All yeast strains employed in this study are described in Table 3.

Standard genetic methods were employed for all procedures (Guthrie & Fink, 1991).

Plasmids

The genomic region containing SSR2 is in GenBank with accession # gplU33050lSCD8035_1 D8035.1p

Plasmids pJU25 and pJU29 are different high copy number suppressors of the temperature sensitive phenotype of *ssr2-1* cloned as random Sau3a1 fragments into the BamH1 site of YEP24. Plasmids pJU134 and pJU135 were derived from pJU25 and pJU29 respectively as Sma1-Sal1 fragments cloned into the polylinker of Bluescript (Stratagene). The Sma1 and Sal1 sites are in regions of YEP24 flanking the inserts. Plasmids pJU163 and pJU164 were derived from pJU134 and pJU135 by removing the Sac1/Sal1 fragment from the Bluescript clone and ligating to the vector PRS424 (2μ, TRP1) (Sikorski & Hieter, 1989). pJU144 contains the bona fide SSR2 clone as a ~12kb Sau3a1 fragment inserted into the BamHI site of YCP50.

U1 plasmids are described elsewhere (Siliciano & Guthrie, 1988; Siliciano et al., 1991).

Other Methods

RNA preparation and primer extension analysis have been described elsewhere (Frank et al., 1992). In vitro splicing and extract preparation were carried out as previously described (Lin et al., 1985). For staged splcing reactions (Figure 7), a standard splicing reaction with mutant extract(s) was incubated for 10 minutes at 25°. At that time, 300-500 fold molar excess cold competitior RNA was added along with additional mutant extract in standard splicing salts. The incubation was continued for 10 more minutes and the products of splicing were analyzed by standard methods.

Complex gel analysis was carried out directly after splicing by adding an equal volume of Heparin Stop Buffer (60 mM KPO4 pH 7.0, 2.5 mM MgCl2, 3% PEG 8000 (same as splicing PEG), ~.8 mgs/ml heparin) and .4 volumes glycerol dye (.25 mg/ml xylene cyanol, 30% glycerol) to reactions placed on ice. Gels were pre-run at 160 Volts for 1 hour at 4°. After loading samles, the gels were run for 2-4 hours at 160 volts until the dye migrated 2/3 down the gel. Gels were constructed by first dissolving .25 gms agarose in 30 mls H2O. When the dissolved agarose had cooled to ~50-60°, a mixture of 6.9 mls acrylamide (60:1), 1.25 mls 1M Tris-acetate pH 7.6, .1 ml 1M KAc, .1 ml 1M MgAc2, .25 mls 10% APS was added and mixed. 3.125 mls DMAPN was then added and mixed and the gel was poured and allowed to solidify for one hour before running. The running buffer contained 25 mM 1M Tris-acetate pH 7.6, 6 mM KAc, 2 mM MgAc2.

Figure 1. Proposed secondary structure of yeast U1 snRNA (Kretzner et al., 1990). Nucleotides 179-533 comprise the yeast specific domain. The nucleotides deleted in U1 Δ10 are boxed.

Saccharomyces cerevisiae U1 snRNA

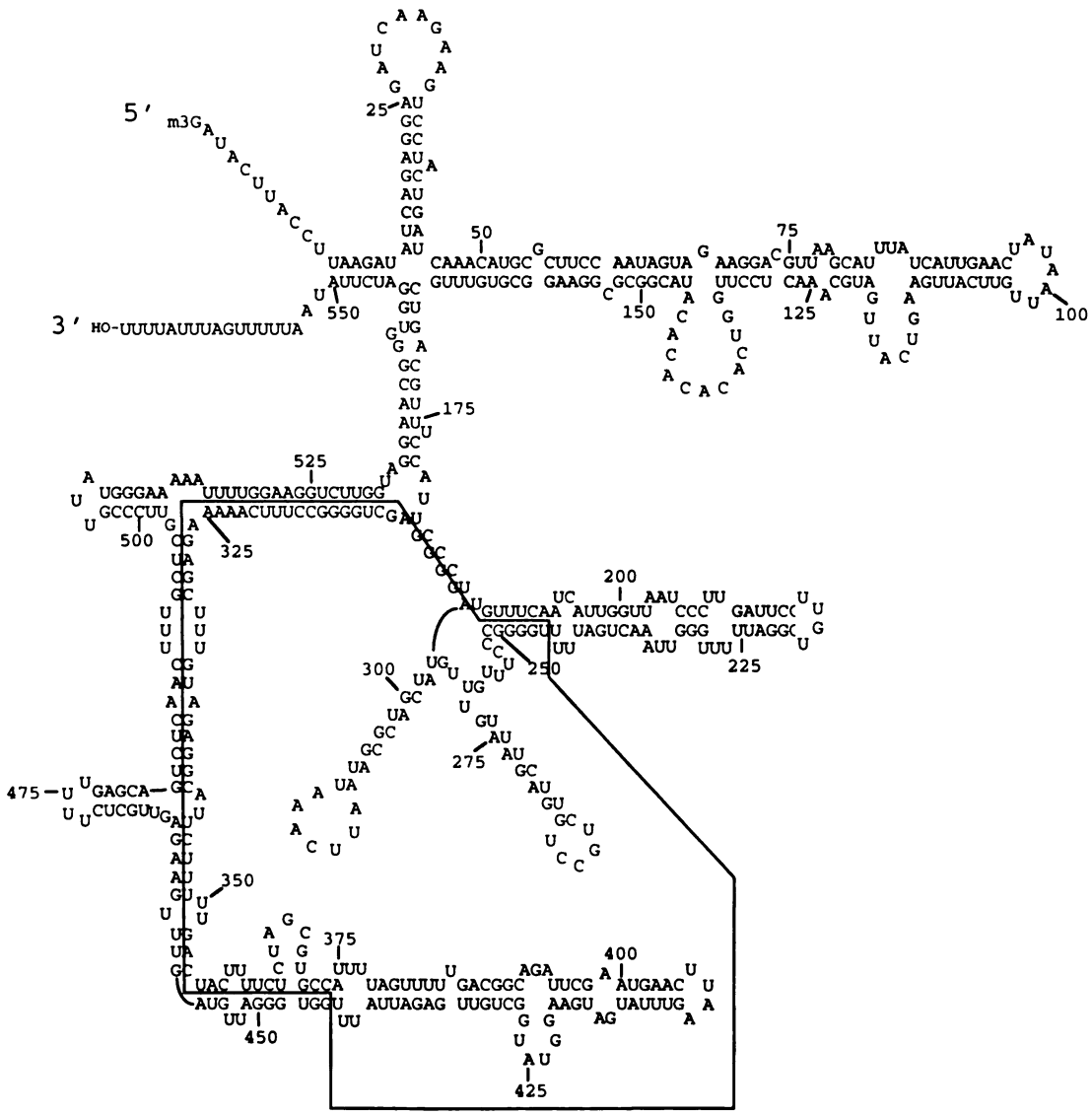


Figure 2. Growth of U1 Δ10 suppressor strains. Each strain is patched in quadruplicate with the genotype given between each panels. Growth was for five days on YEPD at either 34° or 37° as marked.

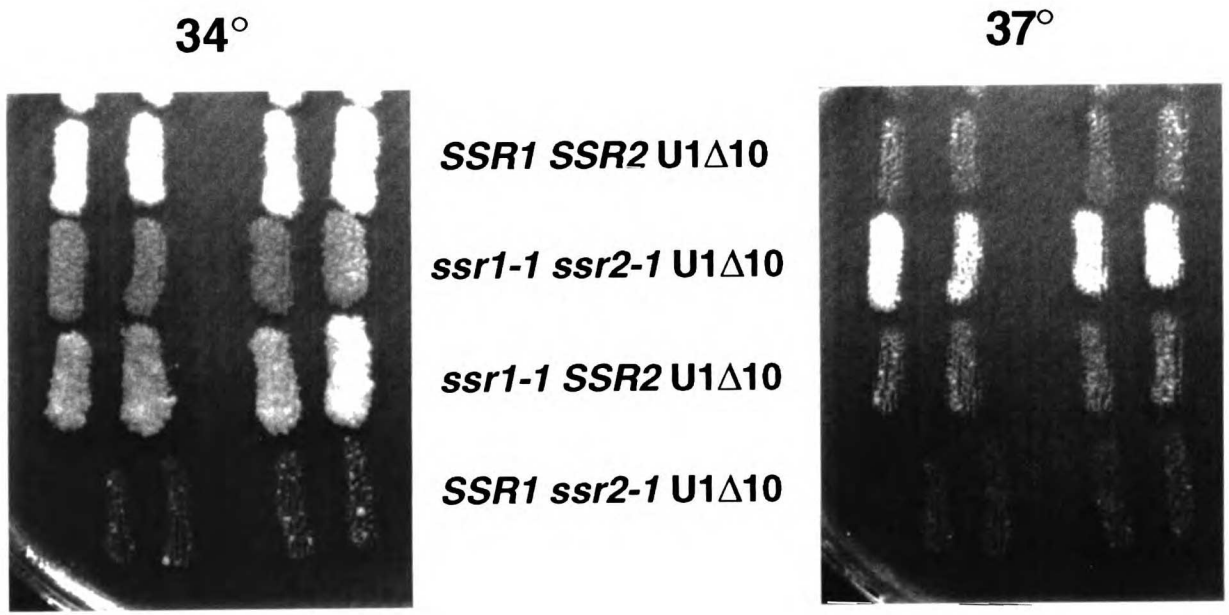

34°
37°
SSR1 SSR2 U1Δ10
ssr1-1 ssr2-1 U1Δ10
ssr1-1 SSR2 U1Δ10
SSR1 ssr2-1 U1Δ10

Figure 3. Unlinked non-complementation between *ssr2-1/slu12-1* and *prp3-1*. Haploid or diploid strains (as marked above the top panel) were constructed and grown at permissive temperature (25°). Ten fold serial dilutions (top to bottom of each lane) were spotted onto complete synthetic media and the plates were incutabed at the temperature indicated next to each panel. The wild type control strain was grown on the same plate but the cells were spotted in a different orientation. Note the poor growth at 37° of the *ssr2-1/prp3-1* and *slu12/prp3-1* diploids.

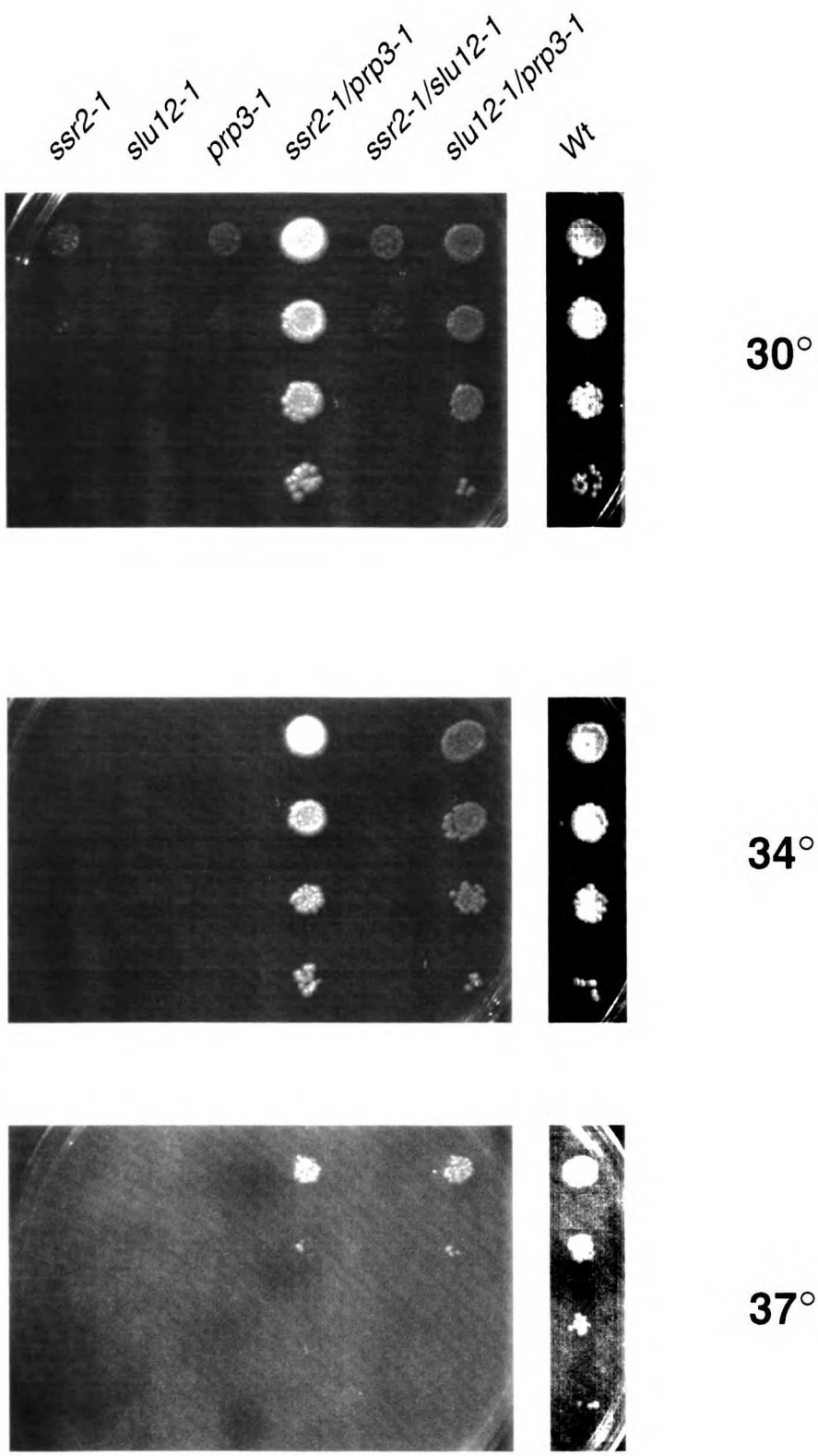

ssr2-1
slu12-1
prp3-1
ssr2-1/prp3-1
ssr2-1/slu12-1
slu12-1/prp3-1
Wt
30°
34°
37°

Figure 4. In vivo splicing phenotype of U1 suppressor strains. Cells were grown for six hours at the indicated temperature, RNA was prepared, and primer extension analysis carried out with primers to rp51a, U1 and U5. Primer extension products are marked to indicate precursor (Pre), lariat intermediate (LI) and mature message (Mat). The genotypes are indicated above each lane. The bottom panel shows a lighter exposure of the region of the gel containing mature. Lanes 1 and 10 (marked M) contain size markers. Note that lane 16 is underloaded.

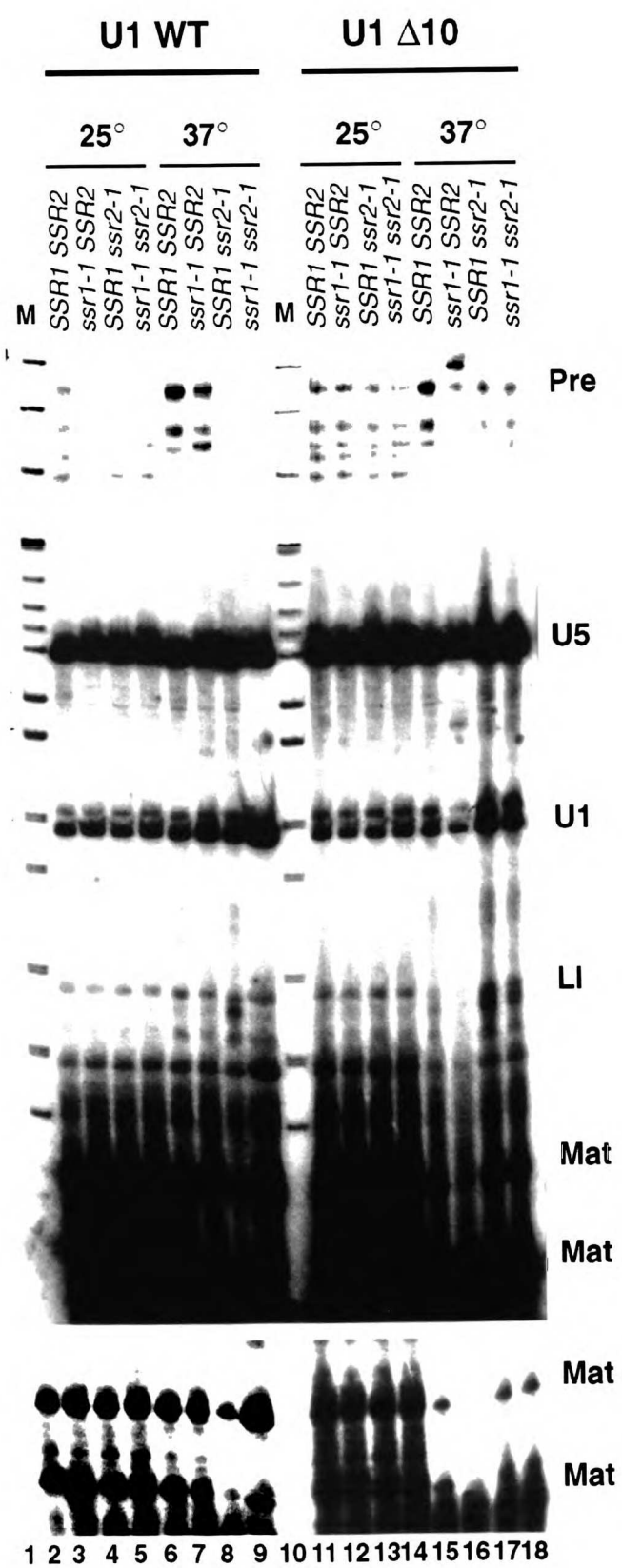

U1 WT
U1 Δ10
25°
37°
25°
37°
M
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
M
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
Pre
U5
U1
LI
Mat
Mat
Mat
Mat
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18

Figure 5. In vitro splicing with U1 suppressor strains. Standard splicing reactions with labeled actin precursor were carried out with extract prepared from the strain indicated above each lane. The positions of lariat intermediate, excised lariat, precursor, mature message and exon 1 (top to bottom) are cartooned next to the gel. Lane 1 contains unreacted precursor.

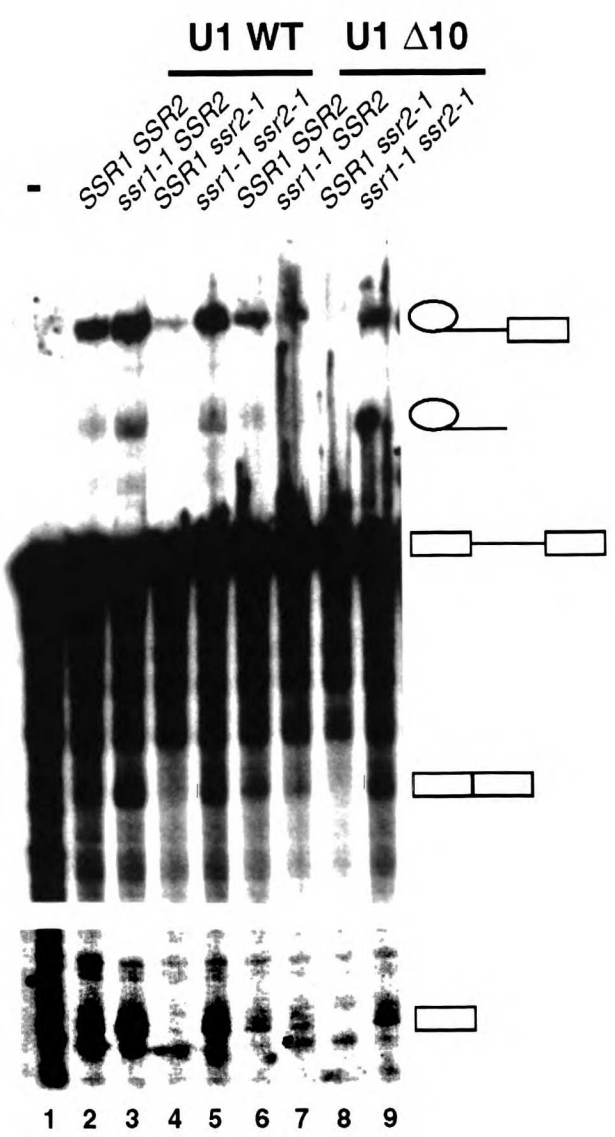

U1 WT
U1 Δ10
-
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
1 2 3 4 5 6 7 8 9

Figure 6. Complex gel analysis of U1 suppressor strains. Reactions from Figure 5 were analyzed on a complex gel. Lane 9 contains a wild type extract incubated without additional ATP. Putative complexes are labeled next to each band and include complete spliceosomes (A1 and A2), U2 containing complexes (B) a novel complex seen only in U1 Δ10 extracts (C) and a non-specific complex.

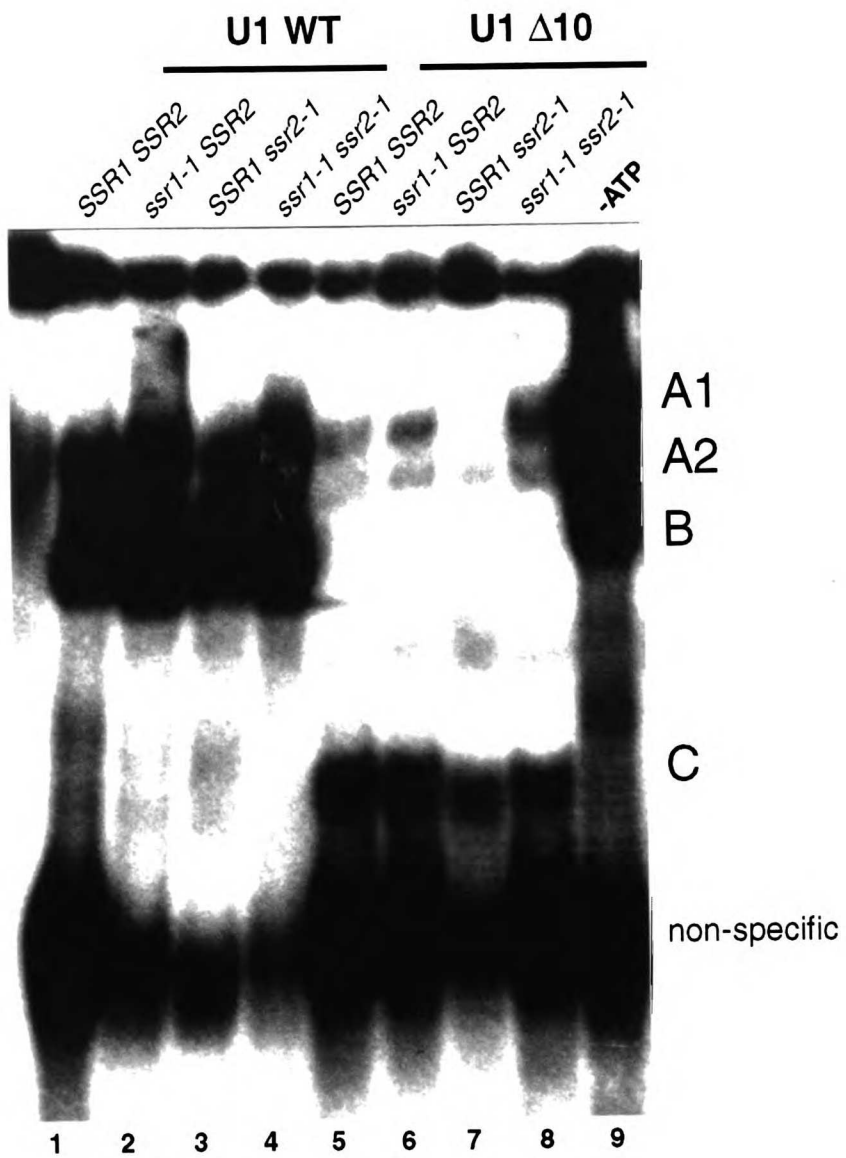

U1 WT
U1 Δ10
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
SSR1 SSR2
ssr1-1 SSR2
SSR1 ssr2-1
ssr1-1 ssr2-1
-ATP
A1
A2
B
C
non-specific
1 2 3 4 5 6 7 8 9

Figure 7. Staged splicing reactions with *ssr2-1* and U1 Δ10 extracts. Splicing reactions were carried out with two separate incubation periods. The presence or absence of each extract or cold competitor RNA in each incubation is indicated above each lane. The incubation in the First Addition was for 10 minutes at 25° and the incubation in the Second Addition was for 10 minutes at 25°. The products of splicing are labeled as in Figure 5.

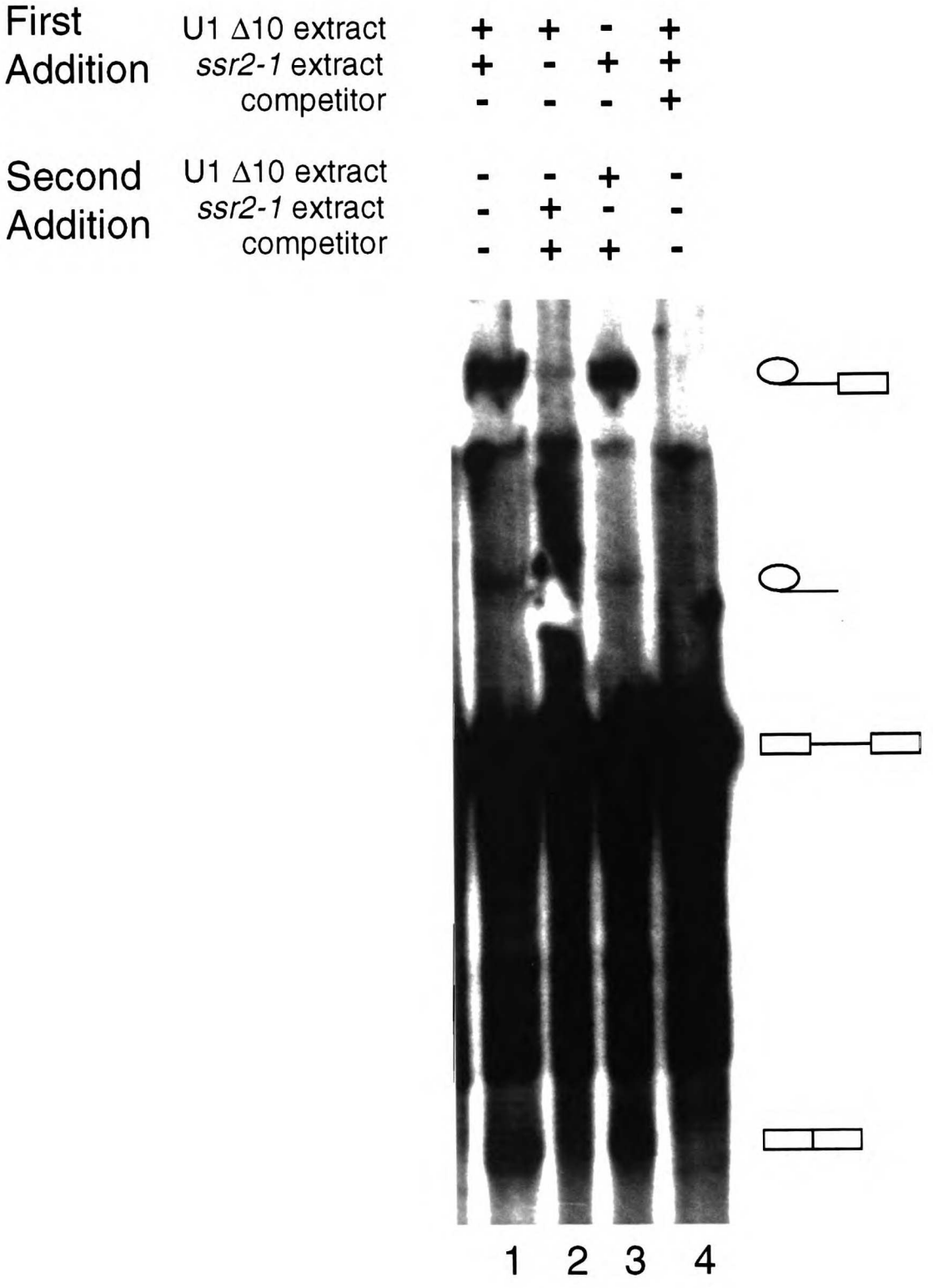

First Addition
U1 Δ10 extract + + - +
ssr2-1 extract + - + +
competitor - - - +
Second Addition
U1 Δ10 extract - - + -
ssr2-1 extract - + - -
competitor - + + -
1 2 3 4

Figure 8. Alignment of C-terminal region of Ssr2 with other HECT-domain-containing proteins. Only the C terminii and preceding region of each family member is shown. The cysteine to which ubiquitin is conjugated is indicated by an asterisk. Residues conserved in all family members, or all but one are in bold. Accession numbers for each protein are: Ssr2, U33050; hs e6ap, L07557; hs 220kd, D28476; hs 115kd, D25215; hs 120kd, D13635; rat 100kd, S22659; sc 168kd, P33202; Rsp5 90kd, L11119. hs, rat and sc stand for human, rat and yeast sequences. Rsp5 is an unpublished yeast protein. This figure was adapted from (Scheffner et al., 1995).

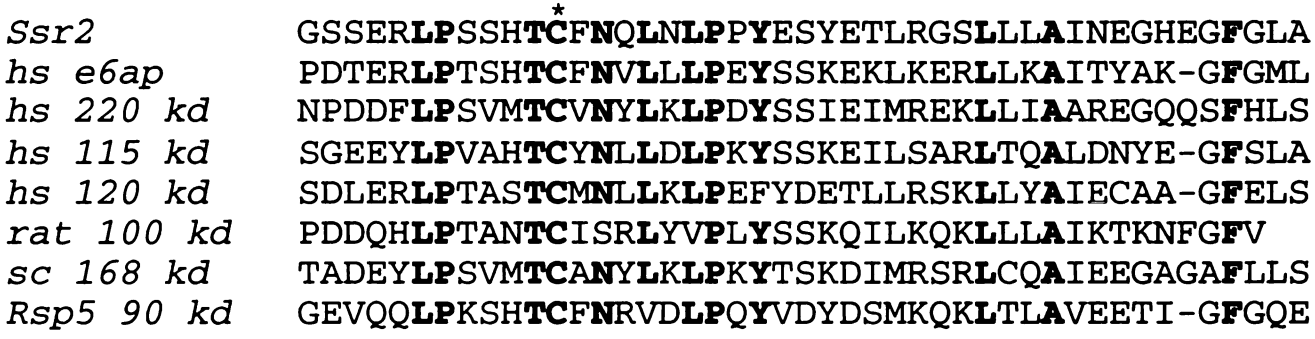

Ssr2 GSSERLPSSHTCFNQLNLPPYESYETLRGSLLLAINEGHEGFGLA
hs e6ap PDTERLPTSHTCFNVLLLPEYSSKEKLKERLLKAITYAK-GFGML
hs 220 kd NPDDFLPSVMTCVNYLKLPDYSSIEIMREKLLIAAREGQQSFHLS
hs 115 kd SGEEYLPVAHTCYNLLDLPKYSSKEILSARLTQALDNYE-GFSLA
hs 120 kd SDLERLPTASTCMNLLKLPEFYDETLLRSKLLYAIECAA-GFELS
rat 100 kd PDDQHLPTANTCISRLYVPLYSSKQILKQKLLLAIKTKNFGFV
sc 168 kd TADEYLPSVMTCANYLKLPKYTSKDIMRSRLCQAIEEGAGAFLLS
Rsp5 90 kd GEVQQLPKSHTCFNRVDLPQYVDYDSMKQKLTLAVEETI-GFGQE

Table 1. Growth of strains containing various U1 (*SNR19*)and suppressor alleles. Growth was scored on YEPD after five days at the indicated temperature. (+++) indicates wild type growth and (-) indicates no growth. All other combinations of + and - symbols represent intermediate levels of growth.

Table 1

Growth phenotypes of U1 suppressor mutations

Strain	Growth Temperature			
	25°	30°	34°	37°
SNR19 SSR1 SSR2	+++	+++	+++	+++
SNR19 ssr1-1 SSR2	+++	+++	+++	+++
SNR19 SSR1 ssr2-1	+++	++	-	-
SNR19 ssr1-1 ssr2-1	+++	++	+	+
snr19 Δ10 *SSR1 SSR2*	++	++	+	-
snr19 Δ10 *ssr1-1 SSR2*	++	++	++	-
snr19 Δ10 *SSR1 ssr2-1*	++	+/-	-	-
snr19 Δ10 *ssr1-1 ssr2-1*	++	++	+	+

Table 2. Allele specificity of suppression. The deleted nucleotides for each U1 allele are indicated in parentheses. Growth is indicated as in Table 1. Instances of suppression are highlighted in boldface.

Table 2

Allele specificity of suppression

Strain	U1 allele	Growth Temperature 18°	34°	37°
SSR1 SSR2	Wt	++	+++	+++
ssr1-1 SSR2	Wt	++	+++	+++
SSR1 ssr2-1	Wt	+	-	-
ssr1-1 ssr2-1	Wt	+	+	+
SSR1 SSR2	Δ10 (246-444)	++	+	-
ssr1-1 SSR2	Δ10 (246-444)	++	++	-
SSR1 ssr2-1	Δ10 (246-444)	+	-	-
ssr1-1 ssr2-1	Δ10 (246-444)	+	+	+
SSR1 SSR2	Δ1 (248-419)	++	+++	-
ssr1-1 SSR2	Δ1 (248-419)	++	+++	-
SSR1 ssr2-1	Δ1 (248-419)	+	-	-
ssr1-1 ssr2-1	Δ1 (248-419)	+	+++	+
SSR1 SSR2	Δ3 (212-418)	++	+	-
ssr1-1 SSR2	Δ3 (212-418)	++	+++	-
SSR1 ssr2-1	Δ3 (212-418)	+	-	-
ssr1-1 ssr2-1	Δ3 (212-418)	+	+++	+
SSR1 SSR2	Δ21 (167-302)	++	-	-
ssr1-1 SSR2	Δ21 (167-302)	++	+/-	-
SSR1 ssr2-1	Δ21 (167-302)	+	-	-
ssr1-1 ssr2-1	Δ21 (167-302)	+	+	-
SSR1 SSR2	Δ27 (156-313)	-	-	-
ssr1-1 SSR2	Δ27 (156-313)	-	-	-
SSR1 ssr2-1	Δ27 (156-313)	+	-	-
ssr1-1 ssr2-1	Δ27 (156-313)	+	-	-
SSR1 SSR2	Δ30 (200-332)	++	+++	-
ssr1-1 SSR2	Δ30 (200-332)	++	+++	-
SSR1 ssr2-1	Δ30 (200-332)	+	-	-
ssr1-1 ssr2-1	Δ30 (200-332)	+	+++	+
SSR1 SSR2	Δ37 (183-313)	++	+++	+/-
ssr1-1 SSR2	Δ37 (183-313)	++	+++	+
SSR1 ssr2-1	Δ37 (183-313)	+	-	-
ssr1-1 ssr2-1	Δ37 (183-313)	+	+++	+

		25°	30°
SSR1 SSR2	U5A	++	nd
ssr1-1 SSR2	U5A	++	nd
SSR1 ssr2-1	U5A	+/-	nd
ssr1-1 ssr2-1	U5A	+/-	nd
SSR1 SSR2	C8U	-	+
ssr1-1 SSR2	C8U	-	+/-
SSR1 ssr2-1	C8U	-	-
ssr1-1 ssr2-1	C8U	-	-
SSR1 SSR2	C8G	-	-
ssr1-1 SSR2	C8G	-	-
SSR1 ssr2-1	C8G	-	-
ssr1-1 ssr2-1	C8G	-	-

Table 3. Yeast strains used in this study. Strains YJU3-10 are spore progeny from the original suppressor mutant. strains 46-65 are isogenic outcrossed derivatives of these strains with various mating type and U1 alleles and different U1 plasmid markers.

Table 3

Yeast strains used in this study

STRAIN	GENOTYPE	U1 ALLELE	U1 PLASMID MARKER
TR2	*Mata his3 trp1 lys2 ura3 ade2*		
YJU46	*Matα his3 trp1 lys2 ura3 ade2*		
	snr19::LYS2 pARR (U1-URA3-CEN-ARS)		
	All strains below derived from YJU46		
YJU 1	*Mata SSR1 SSR2*	Δ10	*TRP1*
YJU 2	*Matα SSR1 SSR2*	Δ10	*TRP1*
YJU 3	*Matα SSR1 SSR2*	Δ10	*TRP1*
YJU 4	*Matα ssr1-1 ssr2-1*	Δ10	*TRP1*
YJU 5	*Mata ssr1-1 ssr2-1*	Δ10	*TRP1*
YJU 6	*Mata SSR1 SSR2*	Δ10	*TRP1*
YJU 7	*Mata SSR1 ssr2-1*	Δ10	*TRP1*
YJU 8	*Matα ssr1-1 SSR2*	Δ10	*TRP1*
YJU 9	*Matα SSR1 ssr2-1*	Δ10	*TRP1*
YJU 10	*Mata ssr1-1 SSR2*	Δ10	*TRP1*
YJU 47	*Matα ssr1-1 SSR2*	WT	*URA3*
YJU 48	*Matα SSR1 ssr2-1*	WT	*URA3*
YJU 49	*Matα ssr1-1 ssr2-1*	WT	*URA3*
YJU 50	*Matα SSR1 SSR2*	Δ10	*URA3*
YJU 51	*Matα ssr1-1 SSR2*	Δ10	*URA3*
YJU 52	*Matα SSR1 ssr2-1*	Δ10	*URA3*
YJU 53	*Matα ssr1-1 ssr2-1*	Δ10	*URA3*
YJU 54	*Matα SSR1 SSR2*	WT	*TRP1*
YJU 55	*Matα ssr1-1 SSR2*	WT	*TRP1*
YJU 56	*Matα SSR1 ssr2-1*	WT	*TRP1*
YJU 57	*Matα ssr1-1 ssr2-1*	WT	*TRP1*
YJU 58	*Matα SSR1 SSR2*	Δ10	*TRP1*
YJU 59	*Matα ssr1-1 SSR2*	Δ10	*TRP1*
YJU 60	*Matα SSR1 ssr2-1*	Δ10	*TRP1*
YJU 61	*Matα ssr1-1 ssr2-1*	Δ10	*TRP1*
YJU 62	*Mata SSR1 SSR2*	Δ10	*TRP1*
YJU 63	*Mata ssr1-1 SSR2*	Δ10	*TRP1*
YJU 64	*Mata SSR1 ssr2-1*	Δ10	*TRP1*
YJU 65	*Mata ssr1-1 ssr2-1*	Δ10	*TRP1*

APPENDIX 3

Future Directions

Below, I outline some of the future directions in which this thesis work might be continued. I shall deliberately keep this section brief, both because of time constraints and because I could probably ramble on ad infinitum about possible models and experiments. I assume that the reader of this section has a good idea of what are the important broad issues and goals for studying the second step. Therefore, I will focus only on specific experimental approaches that might lead towards fulfillment of those goals. For convenience, I have divided this appendix into two sections, biochemistry and genetics.

Biochemistry

1. **Crosslinking.** The technique of site-specific crosslinking has opened up a rich world of possible experiments, into which my thesis work has only briefly forayed. The logical extension of the 3' splice site crosslinking I have done is to repeat these experiments with the branch site and 5' splice site. In particular, the dynamics of RNA-RNA (i.e. U2, 3' splice site) and RNA-protein (i.e. Prp8, Prp16 and Slu7) interactions at the branch site are likely to be very informative with respect to those at the 3' splice site.

In addition, U2, U5 and U6 need to be modified with crosslinking agents, reconstituted, and assayed for crosslinking to Prp8, Prp16 and Slu7. Ideally, a large matrix of information could be generated for timing and placement of different RNA-protein and RNA-RNA interactions during the second step of splicing.

A third direction is to put photoactivatable nucleotides, such as 4-thio-uridine, into the PyAG motif of the 3' splice site. The leaky mutant 3' U<u>G</u>G/ would be an ideal candidate for the incorporation of 4-thio-UpG, and 3' U<u>U</u>G/ is another potential candidate. Alternatively, a guanosine- or adenosine-based crosslinking agent may become available and this could be incorporated directly into a wild type (UAG) 3' splice site. The goal would be to find RNA-RNA crosslinks to the PyAG nucleotides in order to

obtain crucially missing information on the spliceosomal binding site(s) for the 3' splice site.

2. **Prp8p.** With the technique of functional U5 snRNP (and U6) reconstitution available, the next logical step would be to reconstitute snRNPs with recombinant Prp8p. Besides using this system to ask questions about snRNP reconstitution, the added Prp8p could be modified with crosslinking groups so that the RNAs and proteins with which Prp8p interacts during the course of a splicing reaction can be identified.

A second goal would be to prepare various truncated and deleted forms of the protein in the hope of generating a dominant negative allele. While such proteins might have problems competing for wild type Prp8p in vivo, in vitro they could be very powerful reagents for building spliceosomes or snRNPs stuck in novel complexes.

Furthermore, the fact that Prp16p can be cut in half and still function (Yan Wang, personal communication), provides a faint possibility that such an approach might be successful with Prp8p. Division of Prp8p into subdomains might also solve the potential problem of insolubility often encountered with large, heterologously expressed proteins.

3. **Prp16p.** One question which I never got around to answering is whether the Prp16p-induced conformational change at the 3' splice site is irreversible. Specifically, one could ask whether ongoing ATP hydrolysis is necessary to maintain protection of the 3' splice site in Beate Schwer's oligo-directed RNase H assay. The experiment would be to remove ATP from a splicing reaction that contains actin C303/305 (after an incubation to allow lariat intermediate formation) and then probe the 3' splice site. If ATP hydrolysis is required to maintain protection, one could try to figure out if Prp16 was the ATPase involved. The addition of antibodies or dominant negative protein might allow this question to be addressed. If ongoing hydrolysis is not required for protection of the 3' splice site, then the Prp16-dependent reaction is likely to be irreversible, at least in

terms of the protection assay, and no conclusion can be reached as to the number of rounds of ATP hydrolysis required to generate protection. However, if protection of the 3' splice site requires the constant presence of ATP, then this result suggests that multiple rounds of ATP hydrolysis are necessary for the second step of splicing. Such a result might reflect a kinetic proofreading mechanism (see Prp16 section of Epilogue).

4. Slu7p. One question that should be addressed is whether Slu7p is required to splice all 3' splice sites or only the 3'L versions. The observation that both 3'S and 3'L splice sites are impaired in the *slu7-ccss* strain suggests that Slu7p is required for both types of splice sites (Frank & Guthrie, 1992). The *slu7-1* may be special allele impaired for only 3'L splice sites. Since immunodepletion conditions have already been worked out, the next step is to challenge a ΔSlu7 extract with both a 3'L (actin) and 3'S (Mata1 ?, U3 ?) substrate and measure the rate of step two. Repetition with a *slu7-1* extract might also be informative since this allele seems to preferentially affect 3'L splice sites. This experiment could also be done in vivo using genetic depletion of Slu7p.

Another approach is to perform a SELEX experiment with wild type Slu7p (and *slu7-1* encoded protein?). Before doing this, it would be worthwhile to carry out simple binding assays with candidate RNA ligands (e.g. U5, 3' splice site U2/U6).

Genetics

1. Identification of additional second step proteins. Are there more proteins required for step two? Besides Prp8p and Prp16p, the other second step proteins have been identified in screens based on their mutability to a temperature sensitive form (as at least one criteria). Moreover, those screens were not saturating. Therefore, there are probably other second step proteins (Ansari & Schwer, 1995).

The potential gain in demanding temperature sensitivity may not be worth the cost of missing genes that cannot mutate to a temperature sensitive form. I strongly

recommend the use of *prp8-101* as a starting allele in a novel mutant hunt for synthetic lethals. This allele of *PRP8* is ideal since it is already known to interact specifically with other second step mutants. Furthermore, rather than use temperature sensitivity or cold sensitivity as the sole criteria for secondary screens, the candidates could be assayed for exacerbation of a leaky 3' splice site mutation in an *ACT1-CUP1* fusion construct. Thus, three different classes of candidate mutants would be chosen: 1. Those that are temperature sensitive (and show a splicing defect). 2. Those that are cold sensitive (and show a splicing defect). 3. Those that exacerbate the splicing of a 3' splice site mutant and are, therefore, copper sensitive. Class 3 allows for mutants that cannot be isolated as temperature or cold sensitive alleles. Class 3 mutants would still have to be tested in vivo or in vitro for a constitutive splicing defect. Fortunately, random or artifactually produced step two defects do not normally occur when splicing extracts are produced. Therefore, even a partial step two defect from a mutant extract is a good indicator that the mutant is interesting and relevant. Finally, the utility of carrying out such a mutant hunt in strains of each mating type cannot be overemphasized. Tens or hundreds of mutant strains might be quickly pared down by reduction into complementation groups as an initial step. The time spent doing this will save a great deal of subsequent effort in carrying out secondary screens.

2. *PRP8*. *PRP8* is likely to provide a genetic entré into many interesting interactions in the spliceosome. A general approach is to use the mutagenized *PRP8* library described in Chapter 3, in concert with pre-existing genetic assays, to find alleles that specifically affect a given process e.g. branchsite recognition, 5' splice site recognition, U4/U6 unwinding. Such alleles might prove useful as biochemical reagents, or they could be used as genetic probes to find more participants in the process of interest (see above).

An alternative is to screen *PRP8* mutants for dominant negative alleles. Although all classes of such mutants will probably be informative, the alleles which block the

second step of splicing may be particularly useful. While step one-specific dominant negatives must be sorted to determine whether the block is pre- or post-assembly (i.e. before or after Prp8 enters the spliceosome), the step two-specific dominant negatives should all be post-assembly. While any dominant negative mutation in Prp8 that blocks the second step of splicing is worth studying, two types might be especially interesting. The first is a mutant that releases exon 1 prior to the second catalytic step (exon binding defect), and the second is a mutant that causes aberrant cleavages around the 3' splice site (active site defect or 3' splice site binding defect). A third possible phenotype might be hydrolysis at the 3' splice site rather than exon ligation. It is not clear how such a mutant could be identified, except that it might also be partially inhibited for step two (hydrolysis or exon ligation). Other dominant negatives can be tested in various biochemical assays such as 3' splice site crosslinking or Prp16p spliceosomal binding. Finally, isolating second site suppressors will help elucidate the nature of the Prp8p dominant negative defect and identify possible interacting factors.

3. *PRP16*. Based on crosslinking data and a model presented in the epilogue of this thesis, Prp16p is predicted to interact with the 3' splice site. Specifically, there may be a class of *PRP16* alleles, distinct from branchsite suppressors, that suppresses the effects of point mutations at the 3' splice site. Such alleles may affect the ATP hydrolysis/RNA binding cycle of Prp16p at a different point than the branchsite suppressors. Such alleles could be isolated by using a scheme similar to that used by Sean Burgess for isolating branchsite suppressors. In this case, an *ACT1-CUP1* reporter construct with a 3' splice site mutation would be used.

4. *SLU7*. Since Slu7p is likely to be at or near the active site, it might be useful to mutagenize the gene and look for suppressors or enhancers of 3' splice site mutations, or

alleles that affect 3' splice site competitions. Additional temperature- or cold-sensitive alleles would also be useful as a starting point for isolating interacting factors.

5. 2-hybrid screens. The large array of genetic interactions for the second step factors suggests that the proteins interact physically. However, these interactions might be transient and difficult to detect in solution. Therefore, directed 2-hybrid experiments might be the best and easiest way of obtaining evidence for second step protein interactions.

REFERENCES

Aebi M, Hornig H, Padgett RA, Reiser J and Weissmann C. 1986. Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. *Cell 47*:555-565.

Anderson GJ, Bach M, Luhrmann R and Beggs JD. 1989. Conservation between yeast and man of a protein associated with U5 small nuclear ribonucleoprotein. *Nature 342*:819-821.

Ansari A and Schwer B. 1995. SLU7 and a Novel Activity, SSF1, Act Subsequent to PRP16 in the Second Step of Yeast Pre-mRNA Splicing. *EMBO J 14*:4001-4009.

Berget SM, Moore C and Sharp PA. 1977. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A 74*:3171-3175.

Boeke JD, Trueheart J, Natsoulis G and Fink GR. 1987. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. *Methods Enzymol 154*:164-175.

Bone R, Fujishige A, Kettner CA and Agard DA. 1991. Structural basis for broad specificity in alpha-lytic protease mutants. *Biochemistry 30*:10388-10398.

Bordonne R, Banroques J, Abelson J and Guthrie C. 1990. Domains of yeast U4 spliceosomal RNA required for PRP4 protein binding, snRNP-snRNP interactions, and pre-mRNA splicing in vivo. *Genes Dev 4*:1185-1196.

Boulanger SC, Belcher SM, Schmidt U, Dib-Hajj SD, Schmidt T and Perlman PS. 1995. Studies of point mutants define three essential paired nucleotides in the domain 5 substructure of a group II intron. *Mol Cell Biol 15*:4479-4488.

Brown JD and Beggs JD. 1992. Roles of PRP8 protein in the assembly of splicing complexes. *EMBO J 11*:3721-3729.

Burgess S, Couto JR and Guthrie C. 1990. A putative ATP binding protein influences the fidelity of branchpoint recognition in yeast splicing. *Cell 60*:705-717.

Burgess SM and Guthrie C. 1993a. Beat the clock: paradigms for NTPases in the maintenance of biological fidelity. *Trends Biochem Sci 18*:381-384.

Burgess SM and Guthrie C. 1993b. A mechanism to enhance mRNA splicing fidelity: the RNA-dependent ATPase Prp16 governs usage of a discard pathway for aberrant lariat intermediates. *Cell 73*:1377-1391.

Buvoli M, Cobianchi F, Biamonti G and Riva S. 1990. Recombinant hnRNP protein A1 and its N-terminal domain show preferential affinity for oligodeoxynucleotides homologous to intron/exon acceptor sites. *Nucleic Acids Res 18*:6595-6600.

Cellini A, Felder E and Rossi JJ. 1986. Yeast pre-messenger RNA splicing efficiency depends on critical spacing requirements between the branch point and 3' splice site. *EMBO 5*:1023-1030.

Chabot B, Black DL, LeMaster DM and Steitz JA. 1985. The 3' splice site of pre-messenger RNA is recognized by a small nuclear ribonucleoprotein. *Science 230*:1344-1349.

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*:5173-5180.

Chanfreau G and Jacquier A. 1994. Catalytic site components common to both splicing steps of a group II intron. *Science 266*:1383-1387.

Chanfreau G and Jacquier A. 1995. A conformational change beween the two chemical steps of group II self-splicing involves a GNRA tetraloop. *Science* in press.

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*:1981-1987.

Chen JH and Lin RJ. 1990. The yeast PRP2 protein, a putative RNA-dependent ATPase, shares extensive sequence homology with two other pre-mRNA splicing factors. *Nucleic Acids Res 18*:6447.

Cheng SC and Abelson J. 1987. Spliceosome assembly in yeast. *Genes Dev 1*:1014-1027.

Chow LT, Gelinas RE, Broker TR and Roberts RJ. 1977. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell 12*:1-8.

Company M, Arenas J and Abelson J. 1991. Requirement of the RNA helicase-like protein PRP22 for release of messenger RNA from spliceosomes. *Nature 349*:487-493.

Cortes JJ, Sontheimer EJ, Seiwer SD and Steitz JA. 1993. Mutations in the conserved loop of human U5 snRNA generate the use of novel cryptic 5' splice sites in vivo. *EMBO J. 12*:5191-5200.

Couto JR, Tamm J, Parker R and Guthrie C. 1987. A trans-acting suppressor restores splicing of a yeast intron with a branch point mutation. *Genes Dev 1*:445-455.

Crispino JD, Blencowe BJ and Sharp PA. 1994. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science 265*:1866-1869.

Datta B and Weiner AM. 1991. Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature 352*:821-824.

Datta B and Weiner AM. 1993. The phylogenetically invariant ACAGAGA and AGC sequences of U6 small nuclear RNA are more tolerant of mutation in human cells than in Saccharomyces cerevisiae. *Mol Cell Biol 13*:5377-5382.

Deirdre A, Scadden J and Smith CW. 1995. Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. *EMBO J 14*:3236-3246.

Deshler JO and Rossi JJ. 1991. Unexpected point mutations activate cryptic 3' splice sites by perturbing a natural secondary structure within a yeast intron. *Genes Dev 5*:1252-1263.

Dib-Hajj SD, Boulanger SC, Hebbar SK, Peebles CL, Franzen JS and Perlman PS. 1993. Domain 5 interacts with domain 6 and influences the second transesterification reaction of group II intron self-splicing. *Nucleic Acids Res 21*:1797-1804.

Dower WJ, Miller JF and Ragsdale CW. 1988. High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Res 16*:6127-6145.

Elledge SJ and Davis RW. 1988. A family of versatile centromeric vectors for use in the sectoring-shuffle mutagenesis assay in *Saccharomyces cerevisiae*. *Gene 70*:303-312.

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*:404-409.

Fabrizio P and Abelson J. 1992. Thiophosphates in yeast U6 snRNA specifically affect pre-mRNA splicing in vitro. *Nucleic Acids Res 20*:3659-3664.

Fouser LA and Friesen JD. 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.

Fouser LA and Friesen JD. 1987. Effects on mRNA splicing of mutations in the 3' region of the Saccharomyces cerevisiae actin intron. *Mol Cell Biol 7*:225-230.

Frank D and Guthrie C. 1992. An essential splicing factor, SLU7, mediates 3' splice site choice in yeast. *Genes Dev 6*:2112-2124.

Frank D, Patterson B and Guthrie C. 1992. Synthetic lethal mutations suggest interactions between U5 small nuclear RNA and four proteins required for the second step of splicing. *Mol Cell Biol 12*:5197-5205.

Frank DN, Roiha H and Guthrie C. 1994. Architecture of the U5 small nuclear RNA. *Mol Cell Biol 14*:2180-2190.

Frendewey D and Keller W. 1985. Stepwise assembly of a pre-mRNA splicing complex requires U-snRNPs and specific intron sequences. *Cell 42*:355-367.

Freyer GA, Arenas J, Perkins KK, Furneaux HM, Pick L, Young B, Roberts RJ and Hurwitz J. 1987. In vitro formation of a lariat structure containing a G2'-5'G linkage. *J Biol Chem 262*:4267-4273.

Furneaux HM, Perkins KK, Freyer GA, Arenas J and Hurwitz J. 1985. Isolation and characterization of two fractions from HeLa cells required for mRNA splicing in vitro. *Proc Natl Acad Sci U S A 82*:4351-4355.

Garcia-Blanco M, Anderson GJ, Beggs J and Sharp PA. 1990. A mammalian protein of 220 kDa binds pre-mRNAs in the spliceosome: a potential homologue of the yeast PRP8 protein. *Proc Natl Acad Sci U S A 87*:3082-3086.

Garcia-Blanco M, Jamison SF and Sharp PA. 1989. Identification and purification of a 62,000-dalton protein that binds specifically to the polypyrimidine tract of introns. *Genes Dev 3*:1874-1886.

Gerke V and Steitz JA. 1986. A protein associated with small nuclear ribonucleoprotein particles recognizes the 3' splice site of premessenger RNA. *Cell 47*:973-984.

Goodall GJ and Filipowicz W. 1989. The AU-rich sequences present in the introns of plant nuclear pre-mRNAs are required for splicing. *Cell 58*:473-483.

Goux-Pelletan M, Libri D, d'Aubenton-Carafa Y, Fiszman M, Brody E and Marie J. 1990. In vitro splicing of mutually exclusive exons from the chicken beta-tropomyosin gene: role of the branch point location and very long pyrimidine stretch. *EMBO J 9*:241-249.

Gozani O, Patton JG and Reed R. 1994. A novel set of spliceosome-associated proteins and the essential splicing factor PSF bind stably to pre-mRNA prior to catalytic step II of the splicing reaction. *EMBO J 13*:3356-3367.

Green MR. 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. *Annu Rev Cell Biol 7*:559-599.

Guthrie C. 1991. Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein. *Science 253*:157-163.

Guthrie C and Fink GR. 1991. *Guide to yeast genetics and molecular biology*. San Diego: Academic Press

Guthrie C and Patterson B. 1988. Spliceosomal snRNAs. *Annu Rev Genet 1988;22:387-419*

Halfter H and Gallwitz D. 1988. Impairment of yeast pre-mRNA splicing by potential secondary structure-forming sequences near the conserved branchpoint sequence. *Nucleic Acids Res 16*:10413-10423.

Hanahan D. 1983. Studies on transformation of Escherichia coli with plasmids. *J Mol Biol 166*:557-580.

Hanna MM. 1989. Photoaffinity cross-linking methods for studying RNA-protein interactions. *Methods Enzymol 180*:383-409.

Helfman DM and Ricci WM. 1989. Branch point selection in alternative splicing of tropomyosin pre-mRNAs. *Nucleic Acids Res 17*:5633-5650.

Hershko A and Ciechanover A. 1992. The ubiquitin system for protein degradation. *Annu Rev Biochem 61*:761-807.

Hodges PE, Jackson SP, Brown JD and Beggs JD. 1995. Extraordinary sequence conservation of the PRP8 Splicing Factor. *YEAST 11*:337-342.

Hoffman CS and Winston F. 1987. A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli. Gene 57*:267-272.

Hopfield JJ. 1974. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci U S A 71*:4135-4139.

Hornig H, Aebi M and Weissmann C. 1986. Effect of mutations at the lariat branch acceptor site on beta-globin pre-mRNA splicing in vitro. *Nature 324*:589-591.

Horowitz DS and Abelson J. 1993a. Stages in the second reaction of pre-mRNA splicing: the final step is ATP independent. *Genes Dev 7*:320-329.

Horowitz DS and Abelson J. 1993b. A U5 small nuclear ribonucleoprotein particle protein involved only in the second step of pre-mRNA splicing in Saccharomyces cerevisiae. *Mol Cell Biol 13*:2959-2970.

Huibregtse JM, Scheffner M, Beaudenon S and Howley PM. 1995. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proc Natl Acad Sci U S A 92*:5249.

Huibregtse JM, Scheffner M and Howley PM. 1993. Cloning and expression of the cDNA for E6-AP, a protein that mediates the interaction of the human papillomavirus E6 oncoprotein with p53. *Mol Cell Biol 13*:775-784.

Ito H, Fukuda Y, Murata K and Kimura A. 1983. Transformation of intact yeast cells treated with alkali cations. *J. Bacteriol. 153*:163-168.

Jackson SP, Lossky M and Beggs JD. 1988. Cloning of the RNA8 gene of Saccharomyces cerevisiae, detection of the RNA8 protein, and demonstration that it is essential for nuclear pre-mRNA splicing. *Mol Cell Biol 8*:1067-1075.

Jacquier A and Jacquesson-Breuleux N. 1991. Splice site selection and role of the lariat in a group II intron. *J Mol Biol 219*:415-428.

Jacquier A and Michel F. 1987. Multiple exon-binding sites in class II self-splicing introns. *Cell 50*:17-29.

Jacquier A and Michel F. 1990. Base-pairing interactions involving the 5' and 3'-terminal nucleotides of group II self-splicing introns. *J Mol Biol 213*:437-447.

Jacquier A, Rodriguez JR and Rosbash M. 1985. A quantitative analysis of the effects of 5' junction and TACTAAC box mutants and mutant combinations on yeast mRNA splicing. *Cell 43*:423-430.

Jamieson DJ, Rahe B, Pringle J and Beggs JD. 1991. A suppressor of a yeast splicing mutation (prp8-1) encodes a putative ATP-dependent RNA helicase. *Nature 349*:715-717.

Jarrell KA, Dietrich RC and Perlman PS. 1988. Group II intron domain 5 facilitates a trans-splicing reaction. *Mol Cell Biol 8*:2361-2366.

Jones SH, Frank DN and Guthrie C. 1995. Characterization and Functional Ordering of Slu7p and Prp17p During the Second Step of pre-mRNA Splicing in Yeast. *Proc Natl Acad Sci U S A* In Press.

Kim SH and Lin RJ. 1993. Pre-mRNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. *Proc Natl Acad Sci U S A 90*:888-892.

Koch JL, Boulanger SC, Dib-Hajj SD, Hebbar SK and Perlman PS. 1992. Group II introns deleted for multiple substructures retain self-splicing activity. *Mol Cell Biol 12*:1950-1958.

Kohrer K and Domdey H. 1991. Preparation of High Molecular Weight RNA. *Methods in Enzymology 194*:398-405.

Krainer AR and Maniatis T. 1985. Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. *Cell 42*:725-736.

Kramer A, Frick M and Keller W. 1987. Separation of multiple components of HeLa cell nuclear extracts required for pre-messenger RNA splicing. *J Biol Chem 262*:17630-17640.

Kretzner L, Krol A and Rosbash M. 1990. Saccharomyces cerevisiae U1 small nuclear RNA secondary structure contains both universal and yeast-specific domains. *Proc Natl Acad Sci U S A 87*:851-855.

Lamm GM, Blencowe BJ, Sproat BS, Iribarren AM, Ryder U and Lamond AI. 1991. Antisense probes containing 2-aminoadenosine allow efficient depletion of U5 snRNP from HeLa splicing extracts. *Nucleic Acids Res 19*:3193-3198.

Lamond AI, Konarska MM and Sharp PA. 1987. A mutational analysis of spliceosome assembly: evidence for splice site collaboration during spliceosome formation. *Genes Dev 1*:532-543.

Langford CJ, Klinz FJ, Donath C and Gallwitz D. 1984. Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast. *Cell 36*:645-653.

Last RL, Maddock JR and Woolford JL, Jr. 1987. Evidence for related functions of the RNA genes of Saccharomyces cerevisiae. *Genetics 117*:619-631.

Lesser CF and Guthrie C. 1993a. Mutational analysis of pre-mRNA splicing in Saccharomyces cerevisiae using a sensitive new reporter gene, CUP1. *Genetics 133*:851-863.

Lesser CF and Guthrie C. 1993b. Mutations in U6 snRNA that alter splice site specificity: implications for the active site. *Science 262*:1982-1988.

Leung DW, Chen E and Goeddel DV. 1989. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique 1*:11-15.

Liao XC, Tang J and Rosbash M. 1993. An enhancer screen identifies a gene that encodes the yeast U1 snRNP A protein: implications for snRNP protein function in pre-mRNA splicing. *Genes Dev 7*:419-428.

Lin RJ, Lustig AJ and Abelson J. 1987. Splicing of yeast nuclear pre-mRNA in vitro requires a functional 40S spliceosome and several extrinsic factors. *Genes Dev 1*:7-18.

Lin RJ, Newman AJ, Cheng SC and Abelson J. 1985. Yeast mRNA splicing in vitro. *J Biol Chem 260*:14780-14792.

Lindsey LA, Crow AJ and Garcia-Blanco MA. 1995. A mammalian activity required for the second step of pre-messenger RNA splicing. *J Biol Chem 270*:13415-13421.

Lossky M, Anderson GJ, Jackson SP and Beggs J. 1987. Identification of a yeast snRNP protein and detection of snRNP-snRNP interactions. *Cell 51*:1019-1026.

Lustig AJ, Lin RJ and Abelson J. 1986. The yeast RNA gene products are essential for mRNA splicing in vitro. *Cell 47*:953-963.

MacMillan AM, Query CC, Allerson CR, Chen S, Verdine GL and Sharp PA. 1994. Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev 8*:3008-3020.

Madhani HD, Bordonne R and Guthrie C. 1990. Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev 4*:2264-2277.

Madhani HD and Guthrie C. 1992. A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell 71*:803-817.

Madhani HD and Guthrie C. 1994a. Dynamic RNA-RNA interactions in the spliceosome. *Annu Rev Genet 28*:1-26.

Madhani HD and Guthrie C. 1994b. Genetic interactions between the yeast RNA helicase homolog Prp16 and spliceosomal snRNAs identify candidate ligands for the Prp16 RNA-dependent ATPase. *Genetics 137*:677-687.

Madhani HD and Guthrie C. 1994c. Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome. *Genes Dev 8*:1071-1086.

Maniatis T, Fritsch EF and Sambrook J. 1982. *Molecular cloning: A laboratory manual*. Cold Spring Harbor: Cold Spring Harbor Laboratory

Maschhoff KL and Padgett RA. 1993. The stereochemical course of the first step of pre-mRNA splicing. *Nucleic Acids Res 21*:5456-5462.

McPheeters DS and Abelson J. 1992. Mutational analysis of the yeast U2 snRNA suggests a structural similarity to the catalytic core of group I introns. *Cell 71*:819-831.

McSwiggen JA and Cech TR. 1989. Stereochemistry of RNA cleavage by the Tetrahymena ribozyme and evidence that the chemical step is not rate-limiting. *Science 244*:679-683.

Mermoud JE, Cohen P and Lamond AI. 1992. Ser/Thr-specific protein phosphatases are required for both catalytic steps of pre-mRNA splicing. *Nucleic Acids Res 20*:5263-5269.

Michel F, Umesono K and Ozeki H. 1989. Comparative and functional anatomy of group II catalytic introns--a review. *Gene 82*:5-30.

Michels WJ, Jr. and Pyle AM. 1995. Conversion of a group II intron into a new multiple-turnover ribozyme that selectively cleaves oligonucleotides: elucidation of reaction mechanism and structure/function relationships. *Biochemistry 34*:2965-2977.

Miller J. 1972. In: eds. *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp

Milligan JF and Uhlenbeck OC. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol 1989;180:51-62*

Moore MJ, Query CC and Sharp PA. 1993. Splicing of Precursors to mRNA by the Spliceosome. In: R.F. Gesteland and J.F. Atkins, eds. *The RNA World*. Plainview, New York: Cold Spring Harbor Laboratory Press. pp 303-358

Moore MJ and Sharp PA. 1992. Site-specific modification of pre-mRNA: the 2'-hydroxyl groups at the splice sites. *Science 256*:992-997.

Moore MJ and Sharp PA. 1993. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature 365*:364-368.

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

Neer EJ, Schmidt CJ, Nambudripad R and Smith TF. 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature 371*:297-300.

Newman A and Norman C. 1991. Mutations in yeast U5 snRNA alter the specificity of 5' splice-site cleavage. *Cell 65*:115-123.

Newman AJ, Lin RJ, Cheng SC and Abelson J. 1985. Molecular consequences of specific intron mutations on yeast mRNA splicing in vivo and in vitro. *Cell 42*:335-344.

Newman AJ and Norman C. 1992. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell 68*:743-754.

Nilsen TW. 1994. RNA-RNA interactions in the spliceosome: unraveling the ties that bind. *Cell 78*:1-4.

Ninio J. 1975. Kinetic amplification of enzyme discrimination. *Biochimie 57*:587-595.

Padgett RA, Konarska MM, Aebi M, Hornig H, Weissmann C and Sharp PA. 1985. Nonconsensus branch-site sequences in the in vitro splicing of transcripts of mutant rabbit beta-globin genes. *Proc Natl Acad Sci U S A 82*:8349-8353.

Padgett RA, Podar M, Boulanger SC and Perlman PS. 1994. The stereochemical course of group II intron self-splicing. *Science 266*:1685-1688.

Parker R and Guthrie C. 1985. A point mutation in the conserved hexanucleotide at a yeast 5' splice junction uncouples recognition, cleavage, and ligation. *Cell 41*:107-118.

Parker R and Patterson B. 1987. Architecture of fungal introns: Implications for spliceosome assembly. In: M. Inouye and B.S. Dudock, eds. *Molecular Biology of RNA: New Perspectives*. San Diego: Academic Press, Inc. pp 451

Parker R and Siliciano PG. 1993. Evidence for an essential non-Watson-Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. *Nature 361*:660-662.

Patterson B. 1989. *Genetic analysis of the U5 snRNA and the 3' splice site in yeast.* Thesis, University of California, San Francisco.

Patterson B and Guthrie C. 1987. An essential yeast snRNA with a U5-like domain is required for splicing in vivo. *Cell 49*:613-624.

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

Patton JG, Porro EB, Galceran J, Tempst P and Nadal GB. 1993. Cloning and characterization of PSF, a novel pre-mRNA splicing factor. *Genes Dev 7*:393-406.

Peebles CL, Zhang M, Perlman PS and Franzen JS. 1995. Catalytically critical nucleotide in domain 5 of a group II intron. *Proc Natl Acad Sci U S A 92*:4422-4426.

Perkins KK, Furneaux HM and Hurwitz J. 1986. RNA splicing products formed with isolated fractions from HeLa cells are associated with fast-sedimenting complexes. *Proc Natl Acad Sci U S A 83*:887-891.

Piccirilli JA, Vyle JS, Caruthers MH and Cech TR. 1993. Metal ion catalysis in the Tetrahymena ribozyme reaction. *Nature 361*:85-88.

Pikielny CW and Rosbash M. 1985. mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell 41*:119-126.

Pinto AL and Steitz JA. 1989. The mammalian analogue of the yeast PRP8 splicing protein is present in the U4/5/6 small nuclear ribonucleoprotein particle and the spliceosome. *Proc Natl Acad Sci U S A 86*:8742-8746.

Query CC, Moore MJ and Sharp PA. 1994. Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev 8*:587-597.

Rajagopal J, Doudna JA and Szostak JW. 1989. Stereochemical course of catalysis by the Tetrahymena ribozyme. *Science 244*:692-694.

Reed R. 1989. The organization of 3' splice-site sequences in mammalian introns. *Genes Dev 3*:2113-2123.

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

Reed R and Maniatis T. 1988. The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes Dev 2*:1268-1276.

Reich CI, VanHoy RW, Porter GL and Wise JA. 1992. Mutations at the 3' splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. *Cell 69*:1159-1169.

Rein A. 1994. Retroviral RNA packaging: a review. *Arch Virol Suppl 9*:513-522.

Ruby SW and Abelson J. 1991. Pre-mRNA splicing in yeast. *Trends Genet 7*:79-85.

Ruby SW, Chang TH and Abelson J. 1993. Four yeast spliceosomal proteins (PRP5, PRP9, PRP11, and PRP21) interact to promote U2 snRNP binding to pre-mRNA. *Genes Dev 7*:1909-1925.

Ruis BL, Kivens WJ and Siliciano PG. 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*:5190-5195.

Ruskin B and Green MR. 1985. Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. *Nature 317*:732-734.

Ruskin B, Greene JM and Green MR. 1985. Cryptic branch point activation allows accurate in vitro splicing of human beta-globin intron mutants. *Cell 41*:833-844.

Ruskin B, Zamore PD and Green MR. 1988. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell 52*:207-219.

Rymond B and Rosbash M. 1992. Yeast Pre-mRNA Splicing. In: E.W. Jones, J.R. Pringle and J.R. Broach, eds. *The Molecular and Cellular Biology of the Yeast Saccharomyces*. Plainview, New York: Cold Spring Harbor Laboratory Press. pp 143-192

Rymond BC and Rosbash M. 1985. Cleavage of 5' splice site and lariat formation are independent of 3' splice site in yeast mRNA splicing. *Nature 317*:735-737.

Rymond BC, Torrey DD and Rosbash M. 1987. A novel role for the 3' region of introns in pre-mRNA splicing of Saccharomyces cerevisiae. *Genes Dev 1*:238-246.

Sawa H, Ohno M, Sakamoto H and Shimura Y. 1988. Requirement of ATP in the second step of the pre-mRNA splicing reaction. *Nucleic Acids Res 16*:3157-3164.

Sawa H and Shimura Y. 1991. Alterations of RNase H sensitivity of the 3' splice site region during the in vitro splicing reaction. *Nucleic Acids Res 19*:3953-3958.

Schatz PJ, Pillus L, Grisafi P, Solomon F and Botstein D. 1986a. Two functional alpha-tubulin genes of the yeast Saccharomyces cerevisiae encode divergent proteins. *Mol Cell Biol 6*:3711-3721.

Schatz PJ, Solomon F and Botstein D. 1986b. Genetically essential and nonessential alpha-tubulin genes specify functionally interchangeable proteins. *Mol Cell Biol 6*:3722-3733.

Scheffner M, Nuber U and Huibregtse JM. 1995. Protein ubiquitination involving an E1-E2-E3 enzyme ubiquitin thioester cascade. *Nature 373*:81-83.

Schiestl RH and Gietz RD. 1989. High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. *Curr Genet 16*:339-346.

Schmelzer C and Muller MW. 1987. Self-splicing of group II introns in vitro: lariat formation and 3' splice site selection in mutant RNAs. *Cell 51*:753-762.

Schmid SR and Linder P. 1992. D-E-A-D protein family of putative RNA helicases. *Mol Microbiol 6*:283-291.

Schwer B and Guthrie C. 1991. PRP16 is an RNA-dependent ATPase that interacts transiently with the spliceosome. *Nature 349*:494-499.

Schwer B and Guthrie C. 1992a. A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis. *EMBO J 11*:5033-5039.

Schwer B and Guthrie C. 1992b. A dominant negative mutation in a spliceosomal ATPase affects ATP hydrolysis but not binding to the spliceosome. *Mol Cell Biol 12*:3540-3547.

Segault V, Will CL, Sproat BS and Luhrmann R. 1995. In vitro reconstitution of mammalian U2 and U5 snRNPs active in splicing: Sm proteins are functionally interchangeable and are essential for the formation of functional U2 and U5 snRNPs. *EMBO J 14*:4014-4021.

Seraphin B, Abovich N and Rosbash M. 1991. Genetic depletion indicates a late role for U5 snRNP during in vitro spliceosome assembly. *Nucleic Acids Res 19*:3857-3860.

Seraphin B and Kandels-Lewis S. 1993. 3' splice site recognition in S. cerevisiae does not require base pairing with U1 snRNA. *Cell 73*:803-812.

Shannon KW and Guthrie C. 1991. Suppressors of a U4 snRNA mutation define a novel U6 snRNP protein with RNA-binding motifs. *Genes Dev 5*:773-785.

Sikorski RS and Hieter P. 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics 122*:19-27.

Siliciano PG and Guthrie C. 1988. 5' splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. *Genes Dev 2*:1258-1267.

Siliciano PG, Kivens WJ and Guthrie C. 1991. More than half of yeast U1 snRNA is dispensable for growth. *Nucleic Acids Res 19*:6367-6372.

Smith CW, Chu TT and Nadal GB. 1993. Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Mol Cell Biol 13*:4939-4952.

Smith CW and Nadal-Ginard B. 1989. Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell 56*:749-758.

Smith CW, Porro EB, Patton JG and Nadal GB. 1989. Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature 342*:243-247.

Sontheimer EJ. 1994. Site-specific RNA crosslinking with 4-thiouridine. *Mol Biol Rep 20*:35-44.

Sontheimer EJ and Steitz JA. 1993. The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science 262*:1989-1996.

Sorger PK, Ammerer, G. and Shore, D. 1989. Identification and purification of sequence-specific DNA-binding proteins. In: T.E. Creighton, eds. *Protein Function: A Practical Approach*. Oxford: IRL Press. pp 199-223.

Stearns T and Botstein D. 1988. Unlinked noncomplementation: isolation of new conditional-lethal mutations in each of the tubulin genes of Saccharomyces cerevisiae. *Genetics 119*:249-260.

Steitz TA and Steitz JA. 1993. A general two-metal-ion mechanism for catalytic RNA. *Proc Natl Acad Sci U S A 90*:6498-6502.

Strathern JN and Higgins DR. 1991. Recovery of plasmids from yeast into Escherichia coli: shuttle vectors. *Methods Enzymol 194*:319-329.

Suh E and Waring RB. 1992. A phosphorothioate at the 3' splice-site inhibits the second splicing step in a group I intron [published erratum appears in Nucleic Acids Res 1993 Feb 25;21(4):1054]. *Nucleic Acids Res 20*:6303-6309.

Sun JS and Manley JL. 1995. A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes Dev 9*:843-854.

Swanson MS and Dreyfuss G. 1988. RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns. *Embo J 7*:3519-3529.

Tarn WY and Steitz JA. 1994. SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes Dev 8*:2704-2717.

Tazi J, Alibert C, Temsamani J, Reveillaud I, Cathala G, Brunel C and Jeanteur P. 1986. A protein that specifically recognizes the 3' splice site of mammalian pre-mRNA introns is associated with a small nuclear ribonucleoprotein. *Cell 47*:755-766.

Tazi J, Daugeron MC, Cathala G, Brunel C and Jeanteur P. 1992. Adenosine phosphorothioates (ATP alpha S and ATP tau S) differentially affect the two steps of mammalian pre-mRNA splicing. *J Biol Chem 267*:4322-4326.

Teigelkamp S, Newman AJ and Beggs JD. 1995a. 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*:2602-2612.

Teigelkamp S, Whittaker E and Beggs JD. 1995b. Interaction of the yeast splicing factor PRP8 with substrate RNA during both steps of splicing. *Nucleic Acids Res 23*:320-326.

Tinoco I. 1993. Appendix I: Structures of base pairs involving at least two hydrogen bonds. In: E.F. Gesteland and J.F. Atkins, eds. *The RNA World*. Plainview, NY: Cold Spring Harbor Laboratory Press. pp 603-607

Tyers M, Tokiwa G, Nash R and Futcher B. 1992. The Cln3-Cdc28 kinase complex of S. cerevisiae is regulated by proteolysis and phosphorylation. *Embo J 11*:1773-1784.

Umen JG and Guthrie C. 1995a. Mutational analysis of PRP8 reveals domains governing the specificity and fidelity of 3' splice site selection. Submitted.

Umen JG and Guthrie C. 1995b. A novel role for a U5 snRNP protein in 3' splice site selection. *Genes Dev 9*:855-868.

Umen JG and Guthrie C. 1995c. 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*:584-597.

Vankan P, McGuigan C and Mattaj IW. 1992. Roles of U4 and U6 snRNAs in the assembly of splicing complexes. *EMBO J 11*:335-343.

Vijayraghavan U and Abelson J. 1990. PRP18, a protein required for the second reaction in pre-mRNA splicing. *Mol Cell Biol 10*:324-332.

Vijayraghavan U, Company M and Abelson J. 1989. Isolation and characterization of pre-mRNA splicing mutants of Saccharomyces cerevisiae. *Genes Dev 3*:1206-1216.

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*:1683-1695.

Weiner AM. 1993. mRNA splicing and autocatalytic introns: distant cousins or the products of chemical determinism? *Cell 72*:161-164.

Wells SE and Ares M, Jr. 1994. Interactions between highly conserved U2 small nuclear RNA structures and Prp5p, Prp9p, Prp11p, and Prp21p proteins are required to ensure integrity of the U2 small nuclear ribonucleoprotein in Saccharomyces cerevisiae. *Mol Cell Biol 14*:6337-6349.

Whittaker E and Beggs JD. 1991. The yeast PRP8 protein interacts directly with pre-mRNA. *Nucleic Acids Res 19*:5483-5489.

Whittaker E, Lossky M and Beggs JD. 1990. Affinity purification of spliceosomes reveals that the precursor RNA processing protein PRP8, a protein in the U5 small nuclear ribonucleoprotein particle, is a component of yeast spliceosomes. *Proc Natl Acad Sci U S A 87*:2216-2219.

Winkelmann G, Bach M and Luhrmann R. 1989. Evidence from complementation assays in vitro that U5 snRNP is required for both steps of mRNA splicing. *EMBO J 8*:3105-3112.

Wolff T, Menssen R, Hammel J and Bindereif A. 1994. Splicing function of mammalian U6 small nuclear RNA: conserved positions in central domain and helix I are essential during the first and second step of pre-mRNA splicing. *Proc Natl Acad Sci U S A 91*:903-907.

Wu JA and Manley JL. 1991. Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature 352*:818-821.

Wyatt JR, Sontheimer EJ and Steitz JA. 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*:2542-2553.

Yean SL and Lin RJ. 1991. U4 small nuclear RNA dissociates from a yeast spliceosome and does not participate in the subsequent splicing reaction. *Mol Cell Biol 11*:5571-5577.

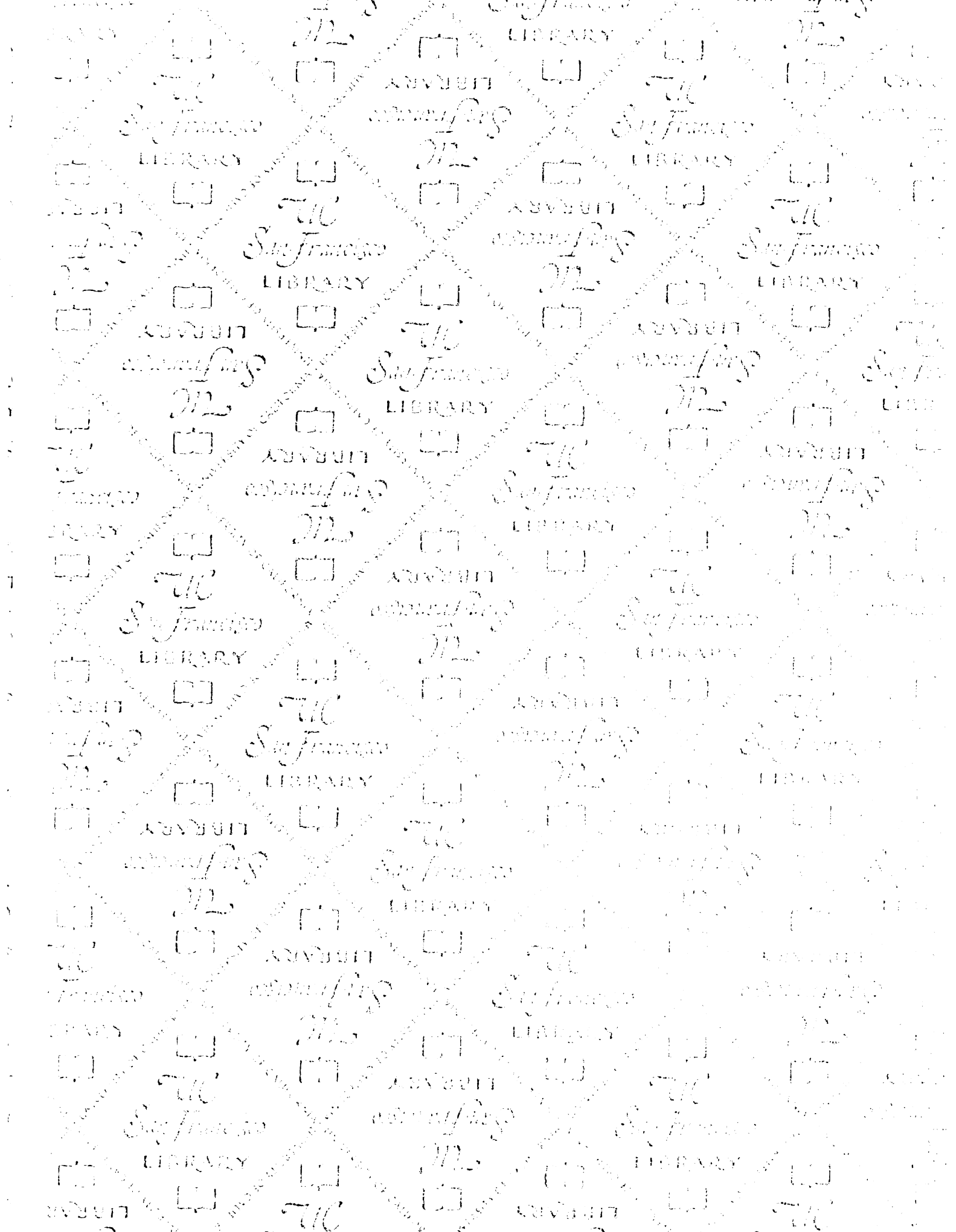
Yisraeli JK and Melton DA. 1989. Synthesis of long, capped transcripts in vitro by SP6 and T7 RNA polymerases. *Methods Enzymol 1989;180:42-50*

Yu YT, Maroney PA, Darzynkiewicz E and Nilsen TW. 1995. U6 snRNA function in nuclear pre-mRNA splicing: A phosphorothioate interference anlaysis of the U6 phosphate backbone. *RNA 1*:46-54.

Yu YT, Maroney PA and Nilsen TW. 1993. Functional reconstitution of U6 snRNA in nematode cis- and trans-splicing: U6 can serve as both a branch acceptor and a 5' exon. *Cell 75*:1049-1059.

Zamore PD and Green MR. 1991. Biochemical characterization of U2 snRNP auxiliary factor: an essential pre-mRNA splicing factor with a novel intranuclear distribution. *Embo J 10*:207-214.

Zhuang Y and Weiner AM. 1990. The conserved dinucleotide AG of the 3' splice site may be recognized twice during in vitro splicing of mammalian mRNA precursors. *Gene 90*:263-269.


San Francisco
LIBRARY