Title
A Genetic analysis of 5' splice site selection in Saccharomyces cerevisiae

Author
Lesser, Cammie F.

Publication Date
1993

Peer reviewed|Thesis/dissertation

A Genetic Analysis of 5' Splice Site Selection in Saccharomyces cerevisiae

by

Cammie F. Lesser

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

[redacted]

Date **University Librarian**

Degree Conferred: 12/31/93

Acknowledgments

I am indebted to many people. First, I wish to thank Christine for her wisdom, for her patience and for teaching me to hang in there and think critically. I thank her for caring about both me and my work and for creating the Guthrie lab, my West Coast home, a very special place that reflects her as a scientist and a person.

I want to thank my thesis committee, Elizabeth Blackburn and Andrew Murray for listening to me and encouraging me as well as offering useful suggestions.

I want to thank Paul Siliciano and Bruce Patterson for turning me on to the awesome power of yeast genetics. Paul for turning me on to the 5' splice site and the question of how the spliceosome knows where to cleave as well as having the good sense not to study this problem in his own lab because it was too risky. I want to thank Bruce Patterson for many late night discussions as well as hot dates to Burger King and Miniature Golf and for his wisdom in suggesting I find a connection between copper resistance and splicing.

I acknowledge that my bay mates, Shelly Jones, Tanya Awabdy and Lucita Esperas, did a great job of putting up with me. I thank Tanya for trying her hardest to keep me humble, but always bringing me treats when she was too mean. Tita for being my surrogate mom and for treating me like one of the boys. Shelly for being an incredible friend and a great reality check. In addition, I thank our adopted bay mate, Pratima Raghunathan, for allowing me to pick her up and turn her over during those especially stressful periods.

I thank my fellow classmate and lab mate, Sean Burgess for always being there to share wacky genetic ideas. Suzanne Noble for being her witty self. Evi Strauss for making up songs about me. Dan Frank for showing me that U5 is Master of the Spliceosome. Jim Umen for venturing into the world of copper along with me. Anita Jandrositz for sharing her lunch. Yan Wong for coming to my championship playoffs. Chris Seibel for being my friend, slave and for taking over softball coaching responsibilities. Hiten Madhani for the U6-U2 interaction. And, Heli for make my life in the lab a lot easier.

I want to thank my family for supporting me through all these years and never forgetting to ask me when would I be finished. I want to thank for instilling in me the curiosity that has taken me this far.

I want to thank Natalie Wu and Marianne Goodman for looking after me and sharing their lives with me. It has been good to have friends on "the outside." Delia Lakich for introducing me to coffee, margaritas and backcountry skiing. As well as coming through for me on those all too many occasions when my life seems to be coming apart at the seams. And lastly, I would like to thank Jon Staley for making sure that I actually did graduate on time. I thank Jon for sharing his enthusiasm over U5 snRNA and the 5' splice site with me as well as for his selflessness, support, friendship and wit.

To: Cammie F. Lesser FAX: (415) 476-0943

Date: December 21, 1993 From: GENETICS Editorial Office

Regular mail: POB 2427, Chapel Hill, NC 27514
Overnight commercial: 108 Porter Place, Chapel Hill, NC 27514
Editor's daytime phone: (919) 541-3361 (John W. Drake)
FAX: (919) 541-7613

You have our permission to include in your dissertation a reprint of your paper in GENETICS **133:** 851–863.

John W. Drake

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

SCHOOL OF MEDICINE
DEPARTMENT OF BIOCHEMISTRY AND BIOPHYSICS
SAN FRANCISCO, CALIFORNIA 94143-0448
(415) 476-4124 FAX # (415) 476-0061

14 December 1993

Editor, Eleanore Butz
Science
1333 H Street NW
Washington, DC 20005

Permission is granted to reprint the material noted in your request. Give credit to SCIENCE. Include in citation: author, title, vol. no., page nos., date of issue, and © AAAS.

12/21/93 (Date)

Arlene Ennis
Permissions Manager
American Association for the Advancement of Science

Dear Dr. Butz,

I am writing to request permission to include in my dissertation a reprint of the paper cited below.

Mutations in U6 snRNA that Alter Splice Site Specificity: Implications for the Active Site, in press.

Please respond by FAX to (415) 476-0943

Thank you for your consideration.

Sincerely,

Cammie Lesser

A Genetic Analysis of 5' Splice Site Selection in *Saccharomyces cerevisiae*

Cammie Faith Lesser

Abstract

A major challenge to the field of pre-mRNA splicing has been to determine how the RNA and protein components of the spliceosome identify splice sites and juxtapose the catalytic residues of the spliceosome with the exon/intron boundaries. U1 snRNA helps to identify the 5' splice site, by interacting directly with the splice site, and to commit the pre-mRNA to the splicing pathway. However, in higher eucaryotes, other factors are involved in the process of 5' splice site recognition. This thesis describes work that implicates additional factors in 5' splice site selection in yeast. In order to screen for factors in both a sensitive and quantitative fashion, a new genetic system was developed in which CUP1 gene fusions are used to monitor pre-mRNA splicing efficiency (Chapter 2). An analysis of several 5' splice site competition assays demonstrates that the process of 5' splice site selection in yeast, like higher eucaryotes, is complicated and involves many determinants including (1) splice site sequence, (2) flanking sequence, (3) proximity to branch point sequence and (4) complementarity to U1 snRNA (Chapters 2, 3 and 4).

Previous work demonstrated that the U1 snRNA·5' splice site interaction is not sufficient to define the exact site of cleavage. This work demonstrates that both U6 and U5 snRNAs

are involved in defining the site of 5' cleavage. U6 snRNA base pairs directly with the 5' splice site and serves to juxtapose the 5' exon-intron boundary with the catalytic residues of the spliceosome (Chapter 5). U5 snRNA can also interact directly with the splice site (Chapter 6). Thus, it appears that U1, U5 and U6 snRNAs are involved in 5' splice site selection. While U1 snRNA appears to commit the pre-mRNA to the splicing pathway and U6 snRNA juxtaposes the 5' exon-intron boundary with the catalytic residues of the spliceosome, the role of U5 snRNA is unclear. Preliminary evidence suggests that U5 snRNA may serve to link commitment and 5' splice site cleavage.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

The interactions of the 5' splice site with snRNAs and proteins

Nuclear pre-mRNA splicing is the process by which introns are removed from mRNA precursors to form functional mature mRNAs. This process is carried out by the spliceosome, a complex ribonucleoprotein machine composed of five small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) and numerous proteins. Introns are removed by two consecutive transesterification reactions (Figure 1). In step 1, the 2'-OH of the intron branch point adenosine attacks the phosphodiester bond at the 5' splice site, resulting in the formation of free exon 1 and lariat intermediate. In step 2, the 3'-OH of exon 1 attacks the 3' splice site resulting in ligation of exon 1 to exon 2 and the release of the excised intron lariat (for review see (Green, 1991; Guthrie, 1991; Moore, et al., 1993)).

The two-step mechanism of nuclear pre-mRNA splicing is similar to the mechanism of intron removal from Group I and Group II self-splicing introns (for review, see (Cech, 1993)). Interestingly, many of the self-splicing introns encode all of the information needed to perform their own excision and can be spliced in the absence of proteins. Each group of self-splicing introns shares a highly conserved secondary structure (Figure 2A). It is the RNA:RNA interactions in these structures, that are often all that is required to juxtapose the cleavage sites with the catalytic residues of the intron and ensure accurate intron excision. In contrast, nuclear pre-mRNA introns are defined solely by the sequence and positioning of three short regions of RNA: the 5' splice site, the branch point region and the 3' splice site (Figure 2B) (for review, see (Green, 1991; Guthrie and Fink, 1991; Moore, et al., 1993)). Thus, a major challenge to the field of pre-mRNA splicing has been to

determine how the RNA and protein components of the spliceosome interact to identify the splice sites and correctly juxtapose the intron-exon boundaries with the catalytic residues of the spliceosome. The major focus of my thesis work has been to understand the role of the 5' splice site in this process.

Roles of the 5' splice site in splicing. Sequence analysis of U1 and U2 snRNAs hinted at their roles in splicing. When a short region of U1 was found to be complementary to the 5' splice site (Figure 3)(Lerner, et al., 1980; Rogers and Wall, 1980), it was proposed that this snRNA could identify the intron by base-pairing with the pre-mRNA (Keller and Noon, 1984). Indeed, genetic experiments, confirmed that U1 and U2 both base pair with the pre-mRNA (Figure 3)(Parker, et al., 1987; Seraphin, et al., 1988; Siliciano and Guthrie, 1988; Wu and Manley, 1989; Zhuang and Weiner, 1986; Zhuang and Weiner, 1989). For example, in yeast, mutation of positions 1, 2 or 5 of the 5' splice site result in a decrease in the efficiency of the first step of splicing, presumably as a result of their impaired ability to base pair with U1. Restoration of Watson-Crick complementarity between U1 and position 1 or 5 of the 5' splice site, by the introduction of compensating mutations in U1, increased the efficiency of step 1 (Seraphin, et al., 1988; Siliciano and Guthrie, 1988). Curiously, neither in yeast nor mammals did the introduction of compensating mutations in U1 restore the level of splicing of the mutated pre-mRNA to wild-type levels. Indeed, in some cases in mammals, the restoration of complementarity between the mutated 5' splice site and U1 had no effect on splicing

(Parker, et al., 1987; Seraphin, et al., 1988; Siliciano and Guthrie, 1988; Wu and Manley, 1989; Zhuang and Weiner, 1986; Zhuang and Weiner, 1989). These observations suggested that recognition of the 5' splice site involves other factors in addition to U1.

Mutational studies of the 5' splice site also revealed that this sequence is involved in several steps of the splicing pathway. For example, all point mutations at intron position 1 (G1), the object of attack of the 2'-OH of the branch point adenosine (see Figure 1), result in a complete block to the second step of splicing but only partially block the first step (Newman, et al., 1985; Fouser and Friesen, 1986; Vijayraghavan, et al., 1986). The second step cannot proceed if the branch point adenosine is covalently bonded to any nucleotide but guanosine. Compensatory changes in U1 do not alleviate the block to the second catalytic step (Siliciano and Guthrie, 1988). This surprising result implies that intron position 1 is examined at least twice in the splicing pathway, first in the pre-mRNA by U1 and later, prior to or during the second catalytic step, in the branched region of the lariat intermediate. Thus, at least one nucleotide of the 5' splice site plays a role in both catalytic steps of splicing.

Intron position 5 also plays at least two roles in the splicing pathway. In *Saccharomyces cerevisiae,* all point mutations at intron position 5 confer the unique phenotype of activating cleavage at an aberrant site in addition to cleavage at the normal 5' splice site (Figure 4A). The location and frequency of aberrant cleavage depends on the particular intron mutated as well as the specific nucleotide change (Jacquier, et al., 1985; Parker and Guthrie, 1985;

Fouser and Friesen, 1986). Interestingly, the lack of sequence complementarity between the new aberrant 5' splice site sequence and U1, argues that this aberrant cleavage site is not defined by a U1·intron base-pairing interaction. In addition, the introduction of compensatory changes in U1 does not prevent aberrant cleavage; rather cleavage is increased at both the normal and aberrant sites (Figure 4B)(Seraphin, et al., 1988; Siliciano and Guthrie, 1988). If base-pairing between U1 and the 5' splice site is an important determinant of cleavage site choice, then the restoration of complementarity between these two RNAs is predicted to prevent cleavage at the aberrant site. Thus, the observed increase in aberrant cleavage indicates that intron position 5 plays at least two roles in the splicing pathway. First, intron position 5 base pairs with U1, as demonstrated by the improved cleavage efficiency seen when U1 is mutated in a manner that restores base-pairing. Second, intron position 5 plays a role in defining the site of 5' cleavage independent of its interaction with U1 snRNA, since compensating U1 mutations do not suppress aberrant cleavage (Seraphin, et al., 1988; Siliciano and Guthrie, 1988).

Previous experiments in the mammalian system had suggested that the base-pairing interaction between U1 and the 5' splice site defined the cleavage site (Aebi, et al., 1987). Consequently, the observation in *S. cerevisiae* that the location of 5' cleavage could be uncoupled from U1 base-pairing was very surprising. Thus, these early experiments raised the question of whether the splicing apparatus of yeast and mammals had evolved different strategies to determine the location of 5' cleavage, an essential step in the

splicing pathway. A related question is whether 5' splice site recognition and cleavage site choice are separate events in mammals?

In summary, mutational studies of the 5' splice site demonstrated that this sequence is involved in several steps of the splicing pathway. These experiments raised the possibility that other factors in addition to U1 interact with this sequence (1) to identify the 5' splice site, (2) to define the exact cleavage site and (3) to determine whether the splicing apparatus should progress through the second step of splicing after inspection of the branched lariat intermediate. Thus, I set out to use genetic techniques to further characterize the roles of the nucleotides of the 5' splice site and to identify factors involved in distinct aspects of intron recognition and/or cleavage site choice.

As mentioned earlier, at the outset of this thesis work U1 snRNA was the only factor known to interact with the 5' splice site. However, during the course of these studies it became clear that this short consensus sequence interacted with additional snRNAs and proteins throughout the splicing pathway.

Interactions between the 5' splice site and snRNAs. U1 is the first snRNA to bind to the pre-mRNA and, in the case of *S. Cerevisiae*, it is the U1·5' splice site interaction that commits the pre-mRNA to the splicing reaction (Legrain, et al., 1988; Seraphin and Rosbash, 1989a). Interestingly, this commitment step is dependent on the presence of both a 5' splice site and a branch point recognition sequence. Presumably, commitment of the pre-mRNA

requires one or more proteins. Only one U1 associated protein, U1-A, has been implicated in commitment complex formation. However, mutation of this protein only prevents commitment of the pre-mRNA if U1 snRNA is also mutated (at a site distinct from the U1-A or 5' splice site binding regions) (Liao, et al., 1993). Mammalian in vitro assays have demonstrated that the U1-C protein helps to stabilize the U1/5' splice site interaction (Heinrichs, et al., 1990); however, no role for this protein in commitment of the pre-mRNA has been demonstrated. To date, no homolog for this protein has been identified in yeast.

U1 and U2 are the only snRNAs that contain obvious stretches of RNA that are complementary to the 5' splice site and branch point recognition consensus sequences. Thus, the likelihood that other snRNAs directly contact the pre-mRNA was presumed to be very low. Consequently, it was very surprising when Newman and Norman (1991) showed that mutations in U5 snRNA could activate an aberrant 5' splice site by base-pairing with the pre-mRNA. They demonstrated that when the 5' splice site was mutated at intron position 1, mutations in the invariant loop of U5 could activate cleavage at aberrant 5' sites. The location of these new cleavage sites was determined by base-pairing interaction between U5 and the sequence just upstream of the cleavage site (positions -2 and -3)(Figure 5)(Newman and Norman, 1991; Newman and Norman, 1992). Notably, however, these special alleles of U5 only activated cleavage at the cryptic sites when the 5' splice site was mutated at position 1. It was thus unclear whether this observation reflected normal U5 function, given that there is no strong conservation of

exon sequence adjacent to normal 5' cleavage site. Interestingly, however, crosslinking studies in mammalian systems demonstrated that U5 is normally in proximity to the exonic region of the 5' splice site (Wyatt, et al., 1992; Sontheimer and Steitz, 1993). It still remains to be seen whether U5 plays an instructive role in normal 5' cleavage site choice.

5' splice site interactions with proteins. Two families of proteins, SR proteins and hnRNP proteins, have been demonstrated to influence 5' splice site selection in higher eucaryotes. All six of the well characterized SR proteins contain both an RNA recognition motif (RRM) and a domain rich in Arg-Ser dipeptides (SR) (Ge and Manley, 1990; Krainer, et al., 1990; Zahler, et al., 1993). Interestingly, the first member of this SR family, ASF/SF-2, was isolated independently as a factor involved in alternative 5' splice site choice (Krainer, et al., 1990) and as an essential splicing factor (Ge and Manley, 1990). Curiously, the addition of any one member of the family to a splicing fraction depleted for all the SR proteins can provide the necessary function to allow splicing to proceed (Krainer, et al., 1990; Zahler, et al., 1993). However, each protein differentially affects the splicing of competing 5' splice sites (Zahler, et al., 1993), suggesting that these proteins may play unique roles in the regulation of alternative splicing.

SR proteins have been shown to commit mammalian pre-mRNAs to the splicing pathway in the absence of U1 (Fu, 1993). This important result argues that a factor interacts with the 5' splice site prior to U1. Interestingly, SR proteins have been demonstrated

to promote U1 binding to the 5' splice site (Zuo and Manley, 1993). No members of the SR family have been found so far in *S. cerevisiae* . If the sole role of these proteins is to regulate alternative splicing, then their absence in *S. cerevisiae* might not be surprising since no examples of alternative splicing have been observed. On the other hand, since, these proteins are essential for mammalian splicing in vitro there may well be yeast homologs.

The mammalian hnRNP protein A1 has also been shown to influence 5' splice site selection (Mayeda and Krainer, 1992). hnRNP proteins are involved in packaging nascent RNAs and presenting them to the RNA processing apparatus. At least some hnRNP proteins have been demonstrated to bind specifically to the polypyrimidine sequence at the 3' splice site of the pre-mRNA (Buvoli, et al., 1988; Gil, et al., 1991; Kumar, et al., 1987; Patton, et al., 1991; Swanson and Dreyfuss, 1988). Interestingly, the crosslinking of two of these proteins to the pre-mRNA, hnRNP A1 and hnRNP C, requires the presence of intact U1 and U2 snRNPs (Mayrand and Pederson, 1990). These observations raise the possibility that interactions between the hnRNPs and snRNPs are important in defining splice site signals. Only recently have hnRNP proteins been found to exist in *S. cerevisiae* (Anderson, et al., 1993), thus it is not yet known if they play a role in yeast pre-mRNA splicing.

Setting up a new genetic system. At the outset, the two goals of this thesis work were: a) to use genetic techniques to further characterize the role of the nucleotides of the 5' splice site; and b) to isolate second site suppressors of the phenotypes conferred by

intron mutations in an attempt to identify trans-acting factors which interact with the 5' splice site. I hoped in this way to address some of the following questions: How are introns recognized? Why does an exon "look" different than an intron? How are the exact cleavage sites chosen?

While second site suppressors of intron mutations are much more difficult to obtain than are loss-of-function (conditional) splicing mutants, one may gain direct insight into function by analyzing the nature of the suppressor mutation. A laborious screen to isolate suppressors of many intron mutations, performed in our laboratory, uncovered only a single mutant allele, prp16-1, a suppressor of an A to C branch point mutation (Couto, et al., 1987). Conditional alleles of PRP16 have been isolated in screens for general splicing mutants; however, it was only through the analysis of branch point suppressor alleles that the role of PRP16 in maintaing the fidelity of splicing was revealed (Burgess, et al., 1990; Burgess and Guthrie, 1993).

One reason why the previous genetic screens had only uncovered a single suppressor of an intron mutation was because of the inadequacies of the genetic reporter systems. Previous genetic screens in the lab had utilized the actin intron fused to either the HIS4 or LacZ (β-galactosidase) reporter genes. Many limitations exist with each of these reporters. First, a high threshold concentration of HIS4 protein needs to be produced in the cell before any growth on media containing histidinol (the biosynthetic precursor of histidine) can be detected (Parker and Guthrie, 1985; Vijayraghavan, et al., 1986). Second, once the threshold level is

obtained, only a small range of increasing amounts of HIS4 fusion protein can be distinguished before the protein is no longer limiting for growth (Vijayraghavan, et al., 1986). Third, when using ACT1-LacZ gene fusions, many unrelated mutations were found to affect the blue color used as an indicator of increased β-galactosidase expression (J. Couto, J. Tamm and C. Guthrie, unpublished). For example, the accumulation of a red pigment in yeast, due to a mutation in the ade1 or ade2 genes, results in an artifactual increase in the blue color of a yeast strain when assessed by plate assays. Finally, LacZ gene fusions can not be used to select suppressors directly; rather the β-galactosidase activity of each mutant candidate must be screened.

Thus, I decided to set up a new reporter system using CUP1, the gene for the yeast metallothionein homolog. CUP1 is normally a non-essential gene, but it allows cells to grow in the presence of copper in a dosage-dependent manner (Hamer, et al., 1985). Previous studies utilizing a CUP1 reporter to assay gene amplification had demonstrated its great sensitivity (Fogel, et al., 1983; Welch, et al., 1983). In addition to the low limit of detection for copper resistance, the system can be used to monitor copper-resistance levels over a large range (Fogel, et al., 1983; Hamer, et al., 1985). Also, by altering the copper levels of media it is possible to select for suppressors directly using CUP1 gene fusions.

Two different strategies were taken to genetically identify factors involved in 5' splice site selection. In addition to searching for second site suppressors of some of the intron mutations discussed earlier, screens were set up to identify factors which

would affect the use of competing 5' splice sites. 5' and 3' splice site competition assays have proven to be a powerful tool for addressing the role of intron and exon sequences in splicing (Eperon, et al., 1986; Fogel, et al., 1983; Hamer, et al., 1985; Nelson and Green, 1988; Patterson and Guthrie, 1991; Reed and Maniatis, 1986; Eperon, et al., 1986; Fogel, et al., 1983; Hamer, et al., 1985; Nelson and Green, 1988; Patterson and Guthrie, 1991; Reed and Maniatis, 1986; J. Umen and C. Guthrie, unpublished). Notably, this type of assay can reveal contributions of components which affect splice site choice but are not normally rate-limiting for splicing. For example, a 3' splice site competition assay was essential for demonstrating a role for a modestly conserved polypyrimidine tract in *S. cerevisiae* (Patterson and Guthrie, 1991).

Chapter 2 describes the construction of the ACT1-CUP1 reporter and a characterization of phenotypes conferred by previously described intron mutations in this new system. In addition this chapter describes the development of a 5' splice site competition assay in the ACT1-CUP1 reporter system in which the 5' splice sites are in close proximity. Chapter 3 describes interactions between U1 and the 5' splice site competition and raises the question of whether yeast has SR-like proteins. Chapter 4 describes two additional 5' splice site competition assays (in which the 5' splice sites are far apart) and discusses the effects of U1 on 5' splice site choice in each of the different assays. Taken together these observations indicate that the process of 5' splice site selection is very similar in both yeast and higher eucaryotes.

snRNA-pre-mRNA interactions: influence on cleavage site choice. The major turning point in my thesis work occurred as a result of a series of in vitro crosslinking experiments in other laboratories. Simultaneously, Wassarman and Steitz (1992) analyzed psoralen induced RNA-RNA crosslinks in mammalian splicing extracts and Sawa and Abelson (1992) characterized ultraviolet (UV) RNA-RNA crosslinks in *S. cerevisiae* splicing extracts. In particular, each group mapped a crosslink between U6 snRNA and the 5' splice site (Figure 6A). Both crosslinks implied a base-pairing interaction between intron position 5 and U6 snRNA. Thus, these experiments raised the exciting possibility that U6 is an additional factor which interacts with the 5' splice site. Specifically, it was possible that disruption of a U6/5' splice site interaction conferred the otherwise inexplicable phenotypes of position 5 intron mutations in activating aberrant splicing. That is, we were intrigued by the exciting possibility that a base-pairing interaction between U6 and the 5' splice site could be a determinant of 5' cleavage site choice. Chapter 5 describes the strategy of the experiments conducted to prove that a base-pairing interaction between U6 and the 5' splice site does indeed occur, and provides evidence that this interaction is an important determinant of the exact location of cleavage (Figure 6B).

Curiously, as discussed earlier, U5 snRNA point mutations can also activate aberrant cleavage sites when the cyh2 intron is mutated at intron position 1 (Newman and Norman, 1991; Newman and Norman, 1992). In Chapter 6, I discuss preliminary experiments that demonstrate that in the case of position 5 mutations in the

actin intron, U5 mutations can increase cleavage at the aberrant site activated by the intron mutation. Thus, in the context of the actin intron, both U5 and U6 mutations have been identified that can alter the amount of cleavage activated at the aberrant site. The epilogue discusses how these snRNAs might be interacting.

Summary. In my attempts to identify factors which interact with the 5' splice site, I began with an unbiased approach using genetic strategies that I hoped would result in the identification of novel components which interact with the 5' splice site. Surprisingly, I eventually identified novel activities of previously characterized splicing factors (indeed, the first identified members of the spliceosome), the snRNAs. However, it was not until the U5 and U6 snRNAs were crosslinked to the pre-mRNA that I realized that these snRNAs might play important roles in cleavage site choice. In retrospect, it is not so surprising that I did not identify factors involved in 5' cleavage site choice using genetics, since it is now apparent that I was looking for a few very specific mutations in very small molecules. Nevertheless, the original genetic strategies resulted in the development of a new highly sensitive CUP1 gene fusion system, which now has many applications. Happily, the crosslinking experiments discussed above provided a very important hint about a role for U5 and U6 in cleavage site choice and allowed me to design experiments which demonstrate that U6 plays an important role in defining the cleavage site. In addition, preliminary evidence indicates that one function of U5 may be to maintain definition of the 5' splice site when U1 leaves and U6 arrives. Thus,

the main contribution of my thesis work has been to characterize the interactions of the 5' splice site with three of five snRNAs (U1, U5 and U6). These studies have resulted in a deeper understanding of how the splicing apparatus recognizes the 5' splice site and chooses the exact cleavage site.

Figure 1. Chemical mechanism of nuclear pre-mRNA splicing. Depicted is the two step mechanism of pre-mRNA splicing. The open boxes represent exons 1 and 2 and are labeled appropriately. The thin black line represents the intron, while the thick black lines with arrows originate from the nucleophilic attacking group and point towards the phosphate bond to be broken. Both of the steps and the splicing intermediates are labeled.

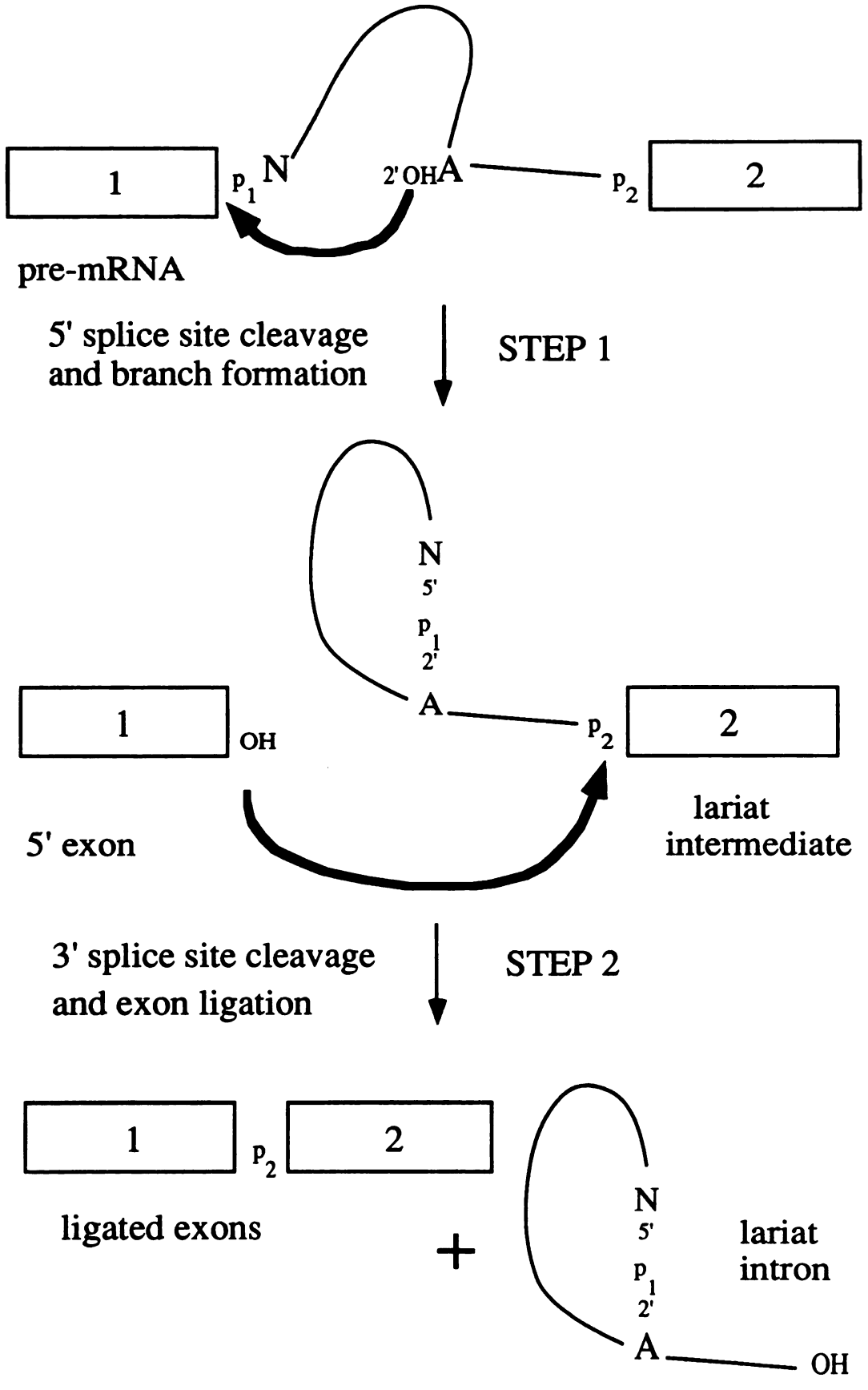
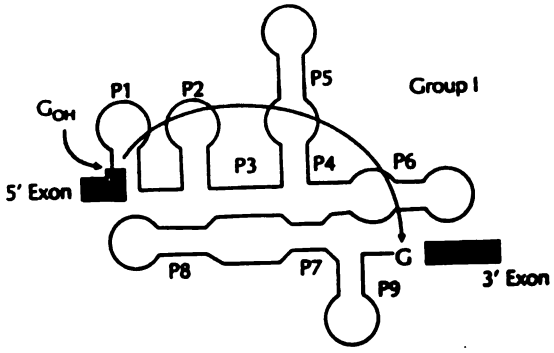
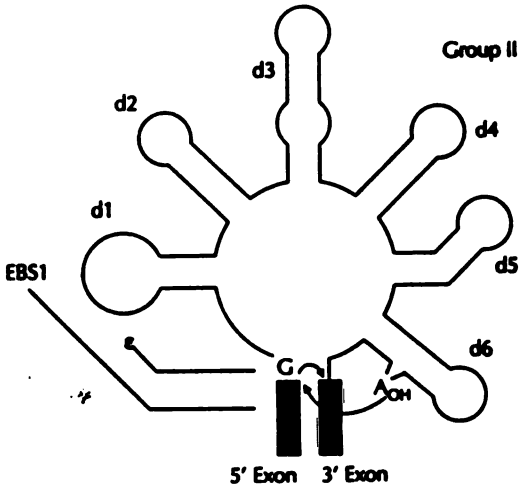

1
p1
N
2'OH
A
p2
2
pre-mRNA
5' splice site cleavage
and branch formation
STEP 1
N
5'
p1
2'
A
1
OH
p2
2
5' exon
lariat
intermediate
3' splice site cleavage
and exon ligation
STEP 2
1
p2
2
ligated exons
+
N
5'
p1
2'
A
OH
lariat
intron

Figure 2 A comparison of Group I, Group II and pre-mRNA intron structures. **(A)**. Depicted are schematic representations of the RNA secondary structure of Group I (top of figure) and Group II (bottom of figure) self-splicing introns. The cleavage sites are represented by arrows. **(B)** Depicted is a schematic representation of a pre-mRNA. Exons 1 and 2 are drawn as open boxes and are labeled appropriately. The intron consensus signals are clearly demarcated.

A.


Group I
P1
P2
P3
P4
P5
P6
P7
P8
P9
G
OH
5' Exon
3' Exon


Group II
d1
d2
d3
d4
d5
d6
EBS1
G
A
OH
5' Exon
3' Exon

B.


5' splice site
branch point
recogntion sequence
3' splice site
1
G U A U G U
U A C U A A C
(PY) n A G
2

Figure 3. U1 and U2 snRNAs base-pair with the 5' splice site and branch point sequence, respectively. Depicted is the base-pairing interaction between U1 and U2 and the 5' splice site and branch point recognition sequence, respectively. Exons 1 and 2 are drawn as open boxes and are labeled appropriately. The intron consensus signals are clearly demarcated.

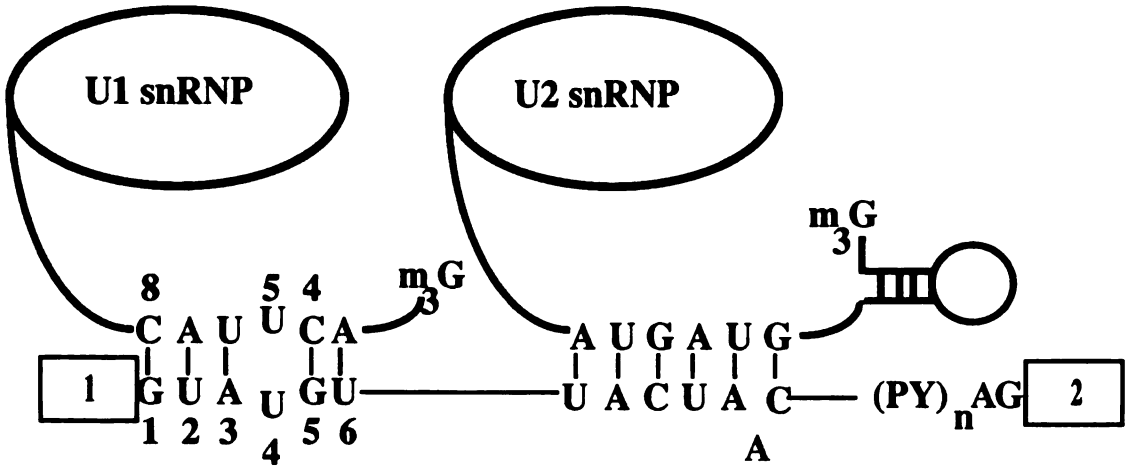
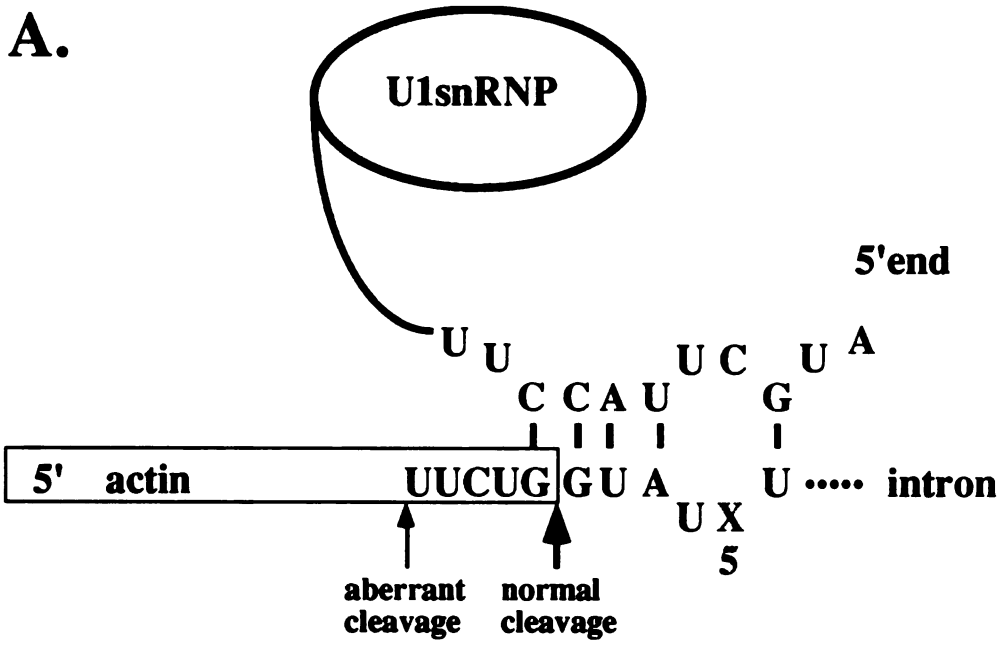

U1 snRNP
U2 snRNP
m3G
m3G
8 5 4
C A U U C A
1 G U A U G U
1 2 3 4 5 6
A U G A U G
U A C U A C
A
(PY)n AG
1
2

Figure 4. 5' splice site mutations activate cleavage at an aberrant site. **(A).** Mutation at actin intron position 5, X, results in the activation of an aberrant cleavage site. **(B).** The restoration of complementarity between U1 and the 5' splice site by the introduction of compensating changes in U1 snRNA, Y, increases cleavage at both the normal and aberrant cleavage sites. The cleavage sites are indicated by arrows. The thickness of the arrows represents the relative amount of cleavage at a particular site.


A.
U1snRNP
5'end
U U U C U A
C C A U G
5' actin UUCUG GU A U ····· intron
U X
5
aberrant cleavage
normal cleavage

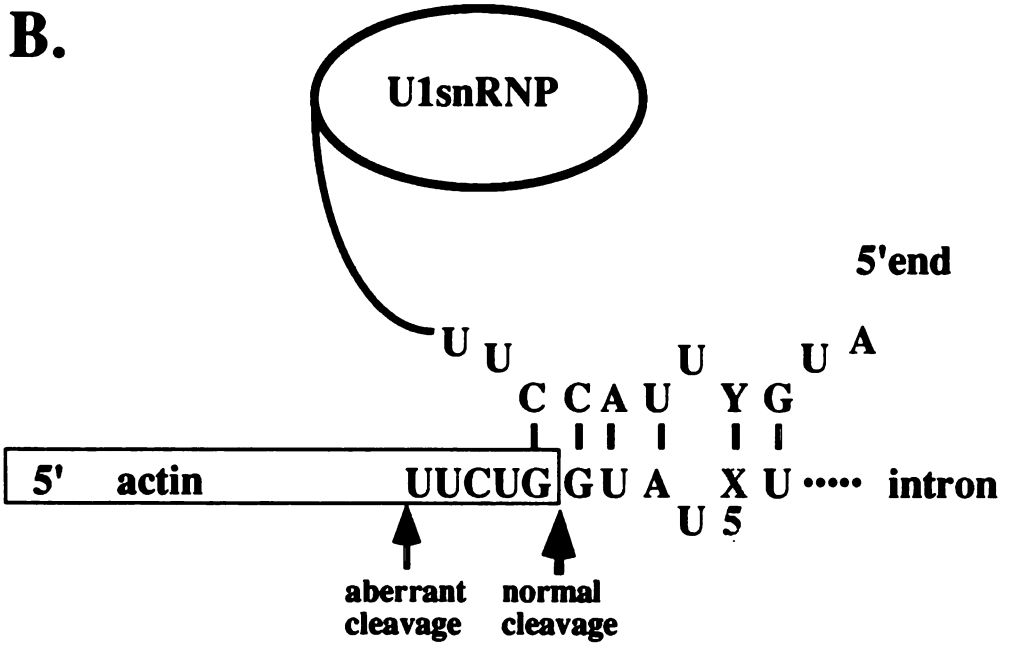

B.
U1snRNP
5'end
U U U U A
C C A U Y G
5' actin UUCUG GU A X U ····· intron
U 5
aberrant cleavage
normal cleavage

Figure 5. U5 snRNA mutants can activate cleavage at aberrant sites. Mutation at cyh2 intron position 1, X, when combined with mutations in the invariant loop of U5 (depicted by open-faced C) can activate aberrant cleavage sites (see text for details).

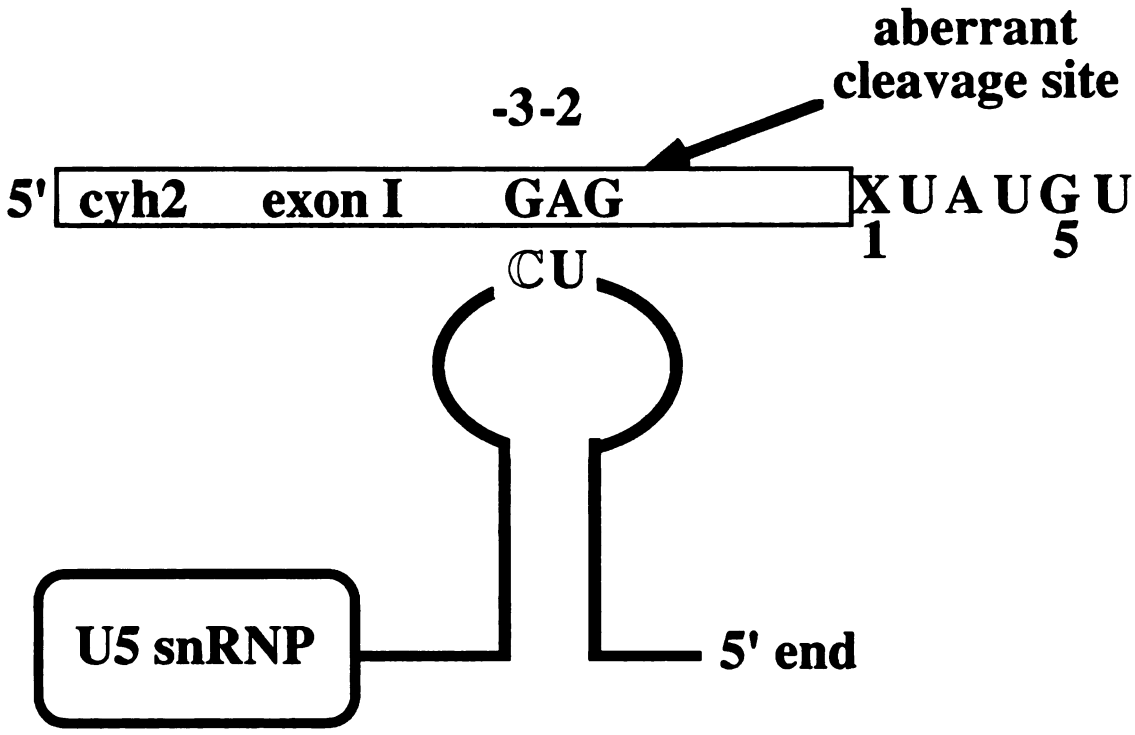
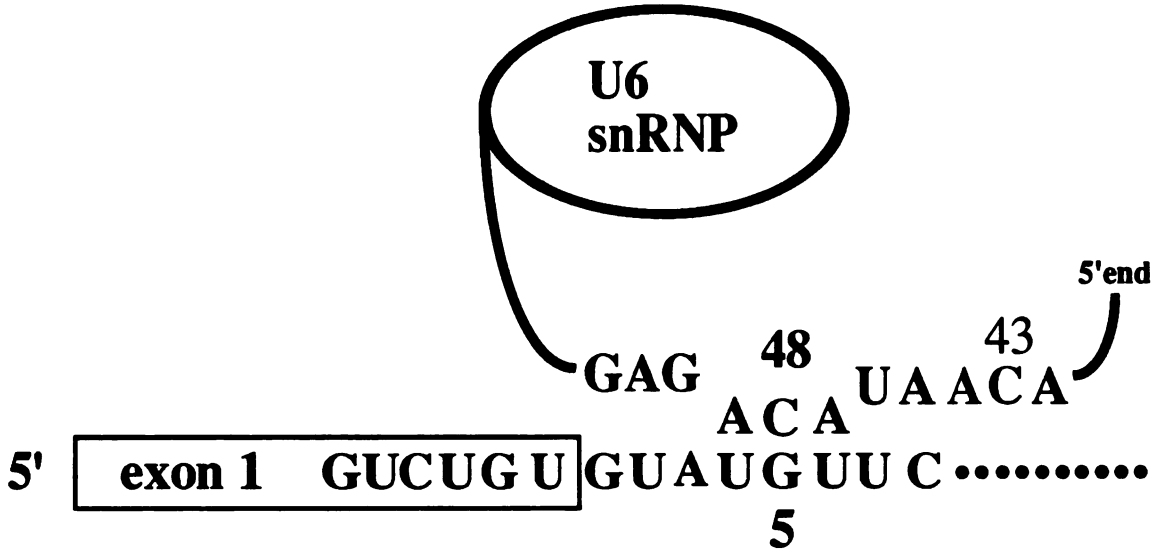
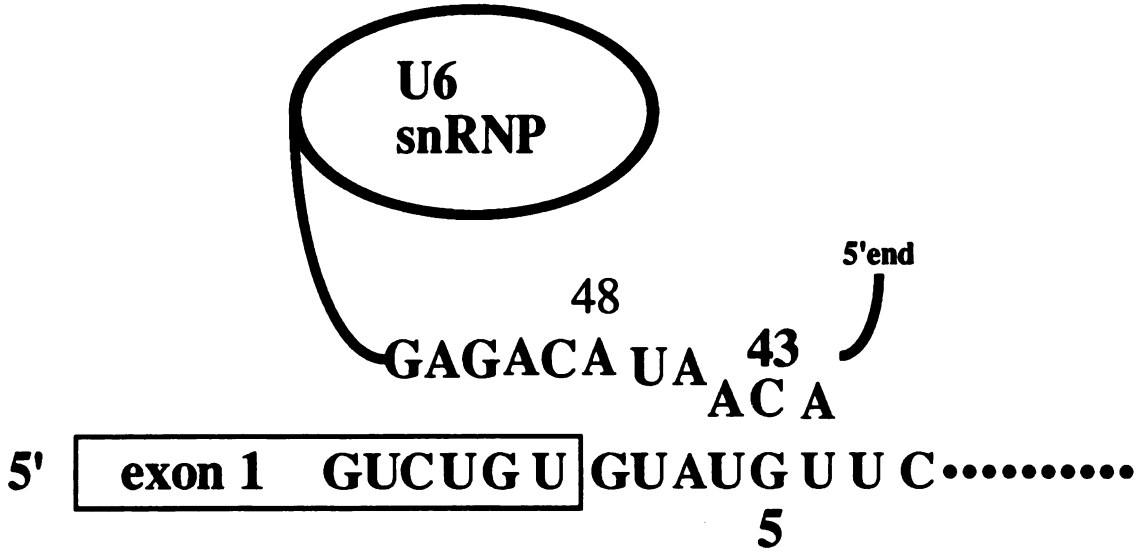

aberrant
cleavage site
-3-2
5'
cyh2
exon I
GAG
X U A U G U
1
5
ℂU
U5 snRNP
5' end

Figure 6. U6 snRNA·5' splice site base-pairing interactions. **(A)** Mammalian psoralen crosslinking studies have resulted in the proposal, depicted on the top of the figure, that U6 snRNA positions 47-49 base-pair with intron positions 4-6 (Wassarman and Steitz, 1992). Yeast ultraviolet crosslinking studies have resulted in the proposal, depicted on the bottom of the figure, that U6 snRNA positions 42-44 base-pair with intron positions 4-6 (Sawa and Abelson, 1992). In both of these models, the invariant and essential ACAGAG sequence of U6 snRNA is at positions 47-52. **(B)** The base-pairing interaction shown between U6 positions 47-49 and intron positions 4-6 does take place and functions to define the cleavage site (marked by the arrow).

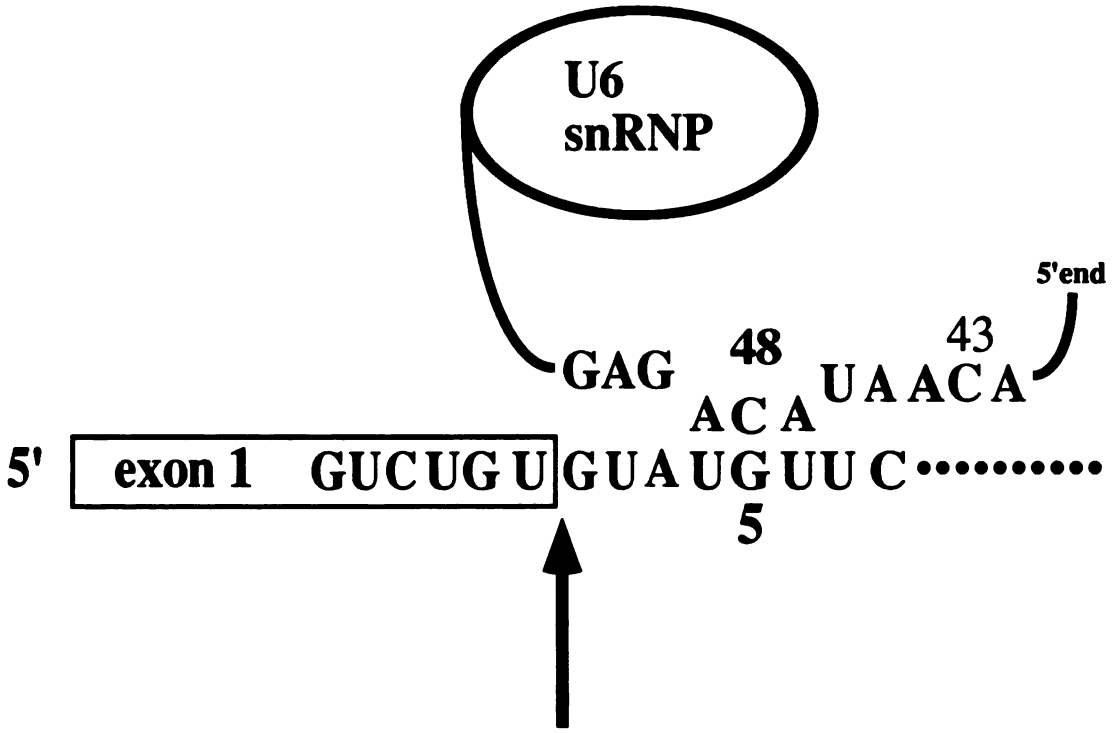


Wassarman and Steitz (1992)



Sawa and Abelson (1992)

NORMAL CLEAVAGE



CHAPTER 2

Mutational Analysis of pre-mRNA splicing in *Saccharomyces cerevisiae* using a sensitive new reporter gene, *CUP1*

ABSTRACT

We have developed a new reporter gene fusion to monitor mRNA splicing in yeast. An intron-containing fragment from the *S.cerevisiae ACT1* gene has been fused to *CUP1*, the yeast metallothionein homolog. *CUP1* is a non-essential gene that allows cells to grow in the presence of copper in a dosage-dependent manner. By inserting previously characterized intron mutations into the fusion construct, we have established that the efficiency of splicing correlates with the level of copper-resistance of these strains. A highly sensitive assay for 5' splice site usage was designed by engineering an *ACT1-CUP1* construct with duplicated 5' splice sites; mutations were introduced into the upstream splice site in order to evaluate the roles of these highly conserved nucleotides in intron recognition. Almost all mutations in the intron portion of the 5' consensus sequence abolish recognition of the mutated site, while mutations in the exon portion of the consensus sequence have variable effects on cleavage at the mutated site. Interestingly, mutations at intron position 4 demonstrate that this nucleotide plays a role in 5' splice site recognition other than by base pairing with U1snRNA. The use of *CUP1* as a reporter gene may be generally applicable for monitoring cellular processes in yeast.

The goal of this work was to develop an improved genetic system by which to monitor pre-mRNA splicing *in vivo*. Intron-containing fragments of several *S. cerevisiae* genes have previously been fused to reporter exons, including *ß-galactosidase (LacZ)* from *E.coli* (Larson, et al., 1983; Teem and Rosbash, 1983), *HIS4* from *S.cerevesiae* (Parker and Guthrie, 1985) and *thymidine kinase* from the Herpes simplex virus (Fouser and Friesen, 1986). These gene fusions have been used (1) to monitor the effects of intron mutations on the production of mature mRNA by measuring the activity of the fusion protein and (2) to genetically identify splicing components involved in intron recognition. To date, there have been only two reports of the successful use of the reporter genes for the genetic identification of trans-acting splicing components (*ACT1-HIS4* and *ACT1-LacZ* gene fusions (Couto, et al., 1987) and *CYH2-LacZ* gene fusions (Newman and Norman, 1991)).

The level of sensitivity and the range of detection of gene fusion reporter products are two very important criteria for the design of useful genetic schemes. *HIS4* gene fusions have limited genetic utility for both of these reasons. First, a high threshold concentration of HIS4 fusion protein needs to be produced in the cell before any growth on media containing histidinol (the biosynthetic precursor of histidine) can be detected (Parker and Guthrie, 1985; (Vijayraghavan, et al., 1986). Second, once the threshold level is obtained, only a small range of increasing amounts of HIS4 fusion protein can be distinguished before the protein is no longer limiting for growth (Vijayraghavan, et al., 1986). In contrast, while *LacZ* fusions are very sensitive and provide relatively linear indicators of

splicing efficiency, these reporters fail a third criterion: they cannot be used for the direct selection of suppressors; rather the ß-galactosidase activity of each mutant candidate must be screened.

In order to avoid many of these limitations, we have designed a new gene fusion using *CUP1* as the reporter. *CUP1* is normally a non-essential gene (Hamer, et al., 1985), but it allows cells to grow in the presence of otherwise lethal concentrations of copper by chelating the metal in a dosage-dependent manner. *CUP1* has been demonstrated to provide a sensitive assay for studying gene amplification (Welch, et al., 1983; Fogel, et al., 1983). The threshold level of detectable CUP1 protein in a cell is very low and there is a large range (0.1-6.0mM $CuSO_4$) over which different copper concentrations can be distinguished by monitoring growth on copper-containing media (Fogel, et al., 1983; Hamer, et al., 1985). In order to facilitate our studies of mRNA splicing, we constructed a gene fusion by combining the 5' exon and intron-containing region of the yeast *ACT1* gene with the complete coding sequence of *CUP1*. In this way we are able to study the efficiency of pre-mRNA splicing by monitoring growth on copper-containing media.

We are interested in determining how introns are recognized. The only common components of all introns are three consensus sequences: the 5' splice site (5'ss), branch point sequence and 3' splice site (3'ss). Mutations in these consensus sequences demonstrate that they are important for both steps of splicing. In step 1, cleavage at the 5' splice site results in the formation of free exon 1 and lariat intermediate. In step 2, ligation of free exon 1 to exon 2 results in the formation of mature mRNA and free intron in

the lariat form (for review, see (Rymond and Rosbash, 1992; Green, 1991; Guthrie, 1991)). Our long-term goal is to employ the *ACT1-CUP1* fusion for the identification of trans-acting splicing factors involved in intron recognition. To evaluate the efficacy of this strategy, we have tested a set of previously characterized mutations in the gene fusion. We find, overall, an excellent correlation between growth on copper and the efficiency of splicing *in vivo.*

We have focused in particular on mutations in the 5' splice site consensus sequence, since this region of the intron is known to have important roles in the earliest steps of spliceosome assembly. In yeast and mammals, the 5' splice site is known to base pair with the 5' end of U1 snRNA, (Figure 1)(Zhuang and Weiner, 1986; Siliciano and Guthrie, 1988; Seraphin, et al., 1988) and, in yeast, the U1 snRNP is the first snRNP to interact with the intron (Ruby and Abelson, 1988; Seraphin and Rosbash, 1989a); if an intact branch point sequence is present, this interaction commits the intron to the splicing pathway (Legrain, et al., 1988; Seraphin and Rosbash, 1989a). Several lines of evidence suggest that the consensus sequences at the 5' splice sites also play other roles in addition to recognition by U1 snRNA. For example, recent genetic experiments in yeast have revealed an interaction between the exon portion of the 5' splice site and U5 snRNA which can influence the precise location of 5' splice site cleavage (Figure 1)(Newman and Norman, 1991). By engineering duplicated 5' splice sites in an *ACT1-CUP1* fusion, we have been able to probe the contribution of each nucleotide in the 5' consensus region to 5' splice site recognition.

Interestingly, our results reveal a role for intron position 4 which does not involve base-pairing with U1 snRNA. Based on these findings, we conclude that the *CUP1* reporter gene fusion, particularly in the context of the splice site competition construct, should provide a sensitive system for the genetic identification of trans-acting components involved in splicing.

MATERIALS AND METHODS

Construction of the wild-type *ACT1-CUP1* plasmid: Site-directed mutagenesis techniques (Kunkel, et al., 1987) were used to introduce a *Sma* I restriction site at the ATG of *CUP1*, such that the ATG of *CUP1* is now GGG(pBS-Cup-Sma). This modification makes it possible to move all of the coding sequence and 167 nucleotides of the 3' untranslated region of *CUP1* (Karin, et al., 1984) on a 353 nucleotide *Sma* I-*Acc* I fragment.

Site-directed mutagenesis techniques (Kunkel, et al., 1987),were next used to introduce a *Hpa* I site 11 nucleotides downstream of the *ACT1* 3' splice site (Ng and Abelson, 1980; Gallwitz and Sures, 1980). An *ACT1-CUP1* fusion was made by cloning the *CUP1 Sma* I-*Acc* I fragment to the modified *ACT1* gene. The sequence at the actin-*CUP1* junction is the following:
3'ss/AGGTTGCTGTT/GGGTTCAGCGAA.

The gene fusion was then removed from the *ACT1* promoter and placed under the control of the glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter on the 2μ pG1 plasmid (Figure 2)(Schena, et al., 1990). Standard cloning techniques (Sambrook, et

al., 1989) as well as PCR methods (Higuchi, 1989) were used to move each of the 5' splice site mutations into the *Bam*H I or *Sal* I site of pG1. In addition, the selectable auxotrophy of pG1, *TRP1*, was replaced by *LEU2* in several clones. The following list specifies the name of the plasmid, the intron mutation, the selectable auxotrophy and the sequence of the junction region between *pG1* and actin. (NOTE: all restriction sites are underlined.)

pGAC1 (WT-*TRP1*):

*GGATCGGTCGAC*CCAATTCGAGCTCGCCCTTTT

pGAC20 (A5-*LEU2*): *GGATCGGTCGAC*CAATTCCTTTT

pGAC30 (U5-*LEU2*): *GGATCC*TTTT

pGAC40 (C5-*LEU2*): *GGATCGGTCGAC*CAATTCGCTTTT

pGAC60 (A1-*LEU2*): *GGATCC*TTTT

pGAC80 (C1-*LEU2*): *GGATCGGTCGAC*CAATTCGCTTTT

pGAC13 was constructed by making a fusion between pGAC1 and pG3 (Schena, et al., 1990). The italic sequence now represents *pG3* sequence.

pGAC13(WT-*TRP1*): *GGATCCCCGGGTACCGAGCTC*GCCCTTTT

pGAC13ΔKpn was made by removing the *Kpn* I site from pGAC13.

pGAC13ΔKpn(WT-*TRP1*): *GGATCCCCGGCGAGCTC*GCCCTTTT

pGAC14 was made by introducing a *Kpn* I (*Asp* 718) near the actin/CUP1 junction so that the sequence there is:

3'ss/AGGTTGGTACC/GGGTTCAGCGA

pGAC24 was made by replacing *TRP1* of pGAC14 with *LEU2*.

pGAC24(WT-*LEU2*): *GGATCCCCGGCGAGCTC*GCCCTTTT

pGAC100(C259-*LEU2*): the C259 mutation was introduced by site-directed mutagenesis techniques (Kunkel, et al., 1987) into pGAC24.

pGACΔ3'ss(*LEU2*) was introduced into pGAC24 by site-directed mutagenesis techniques. The junction sequence between *ACT1* and **CUP1** was altered to be the following:

*TAC*AATATTCATCTC**CGAATTAA**

The sequence at the 3' junction of the gene fusion between *CUP1* and the **pG** vectors is the following in all of the clones:

*TTATAGGTAT*GACCTGCAGCCCAAGCTGGTCGAC

pM3: *prp16-1 (TRP1)* is described in (Burgess, et al., 1990)

pSE358: vector of pM3 (pSE358 was a gift from Dr. S. Elledge)

Construction of the Δcup1 strain: The starting strain was ABDE1 a Δcup1: MAT a, *arg4, CUP1Δ ::URA3, thr1, his4, leu2, ade2, ade5, trp1, ura3, his7*. This strain was provided as a gift from the laboratory of Dr. S Fogel. ABDE1a was backcrossed three times to either TR2a or TR3α (gifts from the laboratory of Dr. P Hieter). This strain was then mated to a yPH399α (gift from Dr. P Hieter) in which we first deleted the *ACT1* intron by an integrative transformation (Rothstein, 1991). The intron deletion is marked by *HIS3*. The *ACT1* intron deletion was confirmed by PCR techniques. The resulting Δcup1 strains are:

K3 aΔcup1: MATa, *cup1Δ::URA3, leu2, ura3, trp1, lys2, ade, GAL+*

I4 αΔcup1: MATα *cup1Δ::URA3, leu2, ura3, trp1 , lys2, ade, GAL+*

Yeast transformations: Yeast lithium acetate transformations were done as described by Becker and Guarente (1991).

Copper growth assays: Copper-resistant growth was determined by plate assays. Copper containing plates were made as described by Sherman (1991); however, Phytagar (Gibco Laboratories) is substituted for Bacto-Agar (Difco). Plates were made to a chosen copper concentration by adding a dilution of filtered $CuSO_4$ (50 or 500mM) solution to the agar solution after autoclaving and before pouring the plates. Copper was always added to complete synthetic media.

Copper resistant growth was determined by either liquid spotting or replica plate assays. Liquid assays were done by growing up saturated overnights of the strain of interest in synthetic media which selects for the *ACT1-CUP1* plasmid. 12.5 microliters of each culture was then micropippeted onto a duplicate set of copper plates of numerous concentrations and growth was assayed after three days in order to determine the level of resistance of the strain. Replica plate assays were done as described by Sherman (1991). Copper-resistant levels were determined three days after replica plating.

RNA preparations: Cells were grown at 30°C in 10 mls of the appropriate selective media to an OD(600) of 0.8-1.2 and then harvested. RNA was then isolated by a scaled down version of the method of Wise (1991). TNE (50mM Tris-HCl, pH 7.4, 100mM NaCl, 10mM EDTA) was substituted for GTE buffer.

Primer extensions: Primer extensions were performed by the method of Frank and Guthrie (1992). The exon 2 primer (3'Cup-

CTTCATTTTGGAAGTTAATTAATT) used is complementary to *CUP1* and lies 11-34 from the *ACT1-CUP1* fusion junction. U1snRNA was primer extended as an internal control for comparing the amount of RNA in each lane. The U1 snRNA primer used is (CAATGACTTCAATGAACAATTAT).

Mutagenesis of the upstream 5' splice site: A 38-nucleotide oligonucleotide was synthesized (see Figure 6) which encodes two 5'splice sites separated by 13 nucleotides. A *Dra* I site is located between the two splice sites. The 5' end of the oligonucleotide encodes a *Bgl* II site while the 3' end of the oligonucleotide encodes a *Nhe* I site. The nine nucleotides of the upstream 5'splice site were synthesized at a low level of degeneracy (97.8% correct, 2.2% mixture of the three incorrect nucleotides). The last six nucleotides of the oligo were designed as a palindrome, to facilitate conversion to a double-strand form through mutually primed synthesis (Oliphant, et al., 1986),. Hybridization and filling in (Oliphant, et al., 1986) yielded a pool of synthetic 5' splice site competition fragments, each one bearing two 5'ss competitions and a central *Nhe* I restriction site. Cleavage with *Bgl* II and *Nhe* I resulted in individual 5'ss competition fragments each with a 5' *Bgl* II end and a 3' *Nhe* I end. 17% of the fragments theoretically should have a single mutation in the upstream 5'ss.

5' Splice Site Competition Plasmid Construction: Site-directed mutagenesis techniques (Kunkel, et al., 1987; Roberts and Zakour, 1987) were used to introduce two restriction sites into

pGAC14. A *Bgl* II site was introduced upstream of the 5'ss and a *Nhe* I site was introduced downstream of the 5'ss. A start codon (ATG) was also engineered just upstream of the *Bgl* II site.

This modified pGAC14 plasmid was digested with *Bgl* II and *Nhe* I and the 5'ss competition fragments were cloned into this vector. The ligation mixture was transformed into bacteria and a library was made by isolating plasmid DNA from 2200 pooled transformants. Since 159 colonies (27 [possible mutations]/.17[probability of having a mutated 5'ss]) includes a complete genome of all possible mutations, the probability that the DNA library contains all possible mutations is 99.99% (Probability $1-e^{-n}$, where n= the number of genomes=13.8).

Characterization of the wild-type competition: Plasmid DNA was isolated from 22 independent bacterial transformants. 91% were of the correct restriction pattern. The wild-type 5'ss competition sequence was identified in 18/20 of the isolated plasmids. The wild-type competition plasmid was then transformed into the Δcup1 strain and analyzed by copper growth and primer extension assays.

The genetic copper selection for 5'splice site competition mutants: The DNA library was transformed into the Δcup1 strain. Several thousand yeast transformants were then replica plated onto 0.25mM copper plates to select for mutants. Library DNA was isolated from approximately 120 transformants which grew on 0.25mM copper by the bead beating method of Strathern and Higgins

(1991) and electroporated into bacteria by the method of Dower, et al. (1988). Plasmids were isolated from bacteria and checked by restriction digest to confirm that only one 5'ss competition fragment had been ligated to the vector (50% of the plasmids had the incorrect restriction pattern). Thirty plasmids which looked correct by restriction map were then sequenced (Table 1) by the method of Sanger, et al. (1977) and retransformed into the Δcup1 strain. Transformants were then assayed for copper-resistance levels. RNA was isolated from these strains and analyzed by primer extensions.

Changing the reading frame of the 5'ss competition plasmids: The 5'ss competition fragments were introduce into pGAC14 so that an *Asp* 718 site lies downstream of the 5' splice sites in the coding sequence. When this restriction site is filled in, the mature RNA, generated as a result of cleavage at the upstream 5'ss, is now in-frame.

Site-directed Mutagenesis: Intron position 6 and exon position -1 were made via dut ung mutagenesis techniques (Kunkel, et al., 1987; Roberts and Zakour, 1987). The exon position -2 mutants were made by PCR-oligo directed mutagenesis (Higuchi, 1989).

Quantitation of Primer Extension Gels: All quantitation was done by scanning with a Molecular Dynamics phosphoimager. Final values were derived by averaging triplicate scans.

RESULTS

Construction of the *ACT1-CUP1* gene fusion: In order to monitor pre-mRNA splicing by assaying levels of copper-resistant growth, we cloned *CUP1*, the *S. cerevisiae* metallothionein homolog (Karin, et al., 1984; Weser, et al., 1977), downstream from an intron-containing fragment of *ACT1*. To maximize the range over which to measure copper-resistant growth, the *ACT1* promoter was replaced by the strong constitutive glyceraldehyde-3-phosphate dehydrogenase (GPD) promoter and the gene fusion was placed on a high copy 2μ vector (Schena, et al., 1990). Thus the fusion gene contains most of exon 1, all of the intron and 11 nucleotides of exon 2 of *ACT1* (Ng and Abelson, 1980; Gallwitz and Sures, 1980) fused to the entire coding sequence of *CUP1* (Figure 2). The start codon of the gene fusion is that of *ACT1*, located ten nucleotides upstream of the intron.

A strain deleted for the chromosomal copy of *CUP1*, Δcup1, is able to grow on media that lack copper, but fails to grow on media that contain >0.013mM copper. Yeast transformed with the wild-type intron-containing gene fusion are able to grow on media that contain 1.0mM copper (Figure 3). Thus, this system provides approximately a 100 fold range in which to measure copper-resistant growth.

Primer extension analyses with a primer complementary to the *CUP1* sequence of exon 2 allow us to quantitate the levels precursor mRNA, lariat intermediate and mature mRNA produced. Primer extension analyses of RNA isolated from yeast containing the wild-

type fusion gene demonstrate that the pre-mRNA can proceed through both steps of splicing (Figure 4A).

Characterization of intron mutations: In order to determine the relationship between splicing efficiency and levels of copper-resistant growth, previously characterized *ACT1* intron mutations were subcloned into the fusion gene. Plasmids containing these intron mutations were transformed into the Δcup1 strain. Mutations were made in all three intron consensus sequences : the 5' splice site (5'ss), branch point sequence and the 3' splice site (3'ss).

Mutations in the 5'ss at intron position 1 (G1A and G1C), as well as mutations which delete the 3'ss (Δ3'ss), have previously been demonstrated to result in a block to the second step of splicing; mature product is not detected in these mutants (Newman, et al., 1985); (Vijayraghavan, et al., 1986)Vijayraghavan, et al., 1986; (Fouser and Friesen, 1986). Δcup1 strains containing *ACT1-CUP1* gene fusions with G1, G1C or Δ3'ss mutations grew like strains which contain no *ACT1-CUP1* gene fusions. That is, they were unable to grow on media which contain >0.013mM copper (Figure 3, Table 2).

Primer extension analyses of RNA isolated from these strains demonstrate that G1A, G1C and Δ3'ss are unable to undergo the second step of splicing; they all accumulate lariat intermediate and produce no mature mRNA (Figure 4A). Mutations at intron position 1 also accumulate precursor mRNA as a result of a decrease in the efficiency of step 1. As shown previously (Vijayraghavan, et al.,

1986), this effect is much greater with G1C than G1A. Δ3'ss has only a slight effect on the efficiency of step 1 (Figure 4A).

All three intron position 5 mutations are able to grow on copper plates, in the order: G5C > G5U ≥ G5A (Figure 3, Table 2). These mutations result in a decrease in the efficiency of both steps of splicing (Parker and Guthrie, 1985; Jacquier, et al., 1985; Fouser and Friesen, 1986; Lesser and Guthrie, unpublished). Primer extension analyses with an exon 2 primer demonstrate that precursor, lariat intermediate and mature products are present in each case (Figure 4A). Most importantly, when normalized to the amount of an internal control, the amount of mature product correlates with relative growth on copper plates: C5>U5≥A5. Thus, growth on copper-containing media accurately reflects the amount of spliced *ACT1-CUP1* mRNA.

The last intron mutation examined was the *ACT1* branch point mutation A259C, which has also been shown to result in a decrease in the efficiency of both steps of splicing (Newman, et al., 1985). Precursor, lariat intermediate and mRNA are observed by primer extension analyses (Figure 4B), demonstrating that the blocks to both steps of splicing are not complete. However, these strains produce less mRNA relative to an internal control compared to the strains transformed with intron position 5 mutations. Consistent with these results is the observation that Δcup1 strains containing this plasmid are less copper-resistant than Δcup1 strains containing intron position 5 mutations (Figure 3, Table 2).

Genetic suppressor reconstruction experiment: The above results demonstrate that the phenotypes conferred by intron mutations introduced into the high copy *ACT1-CUP1* gene fusion closely reflect the phenotypes conferred by the same intron mutations on low copy *ACT1-HIS4* and *ACT1*-LacZ fusions. The increased sensitivity of the *ACT1-CUP1* system (see discussion) and the facility of the copper growth assays demonstrate that the *ACT1-CUP1* system is a quick and sensitive assay for characterizing *cis* mutations. To determine if the *ACT1-CUP1* system could also be used to find trans-acting suppressors of intron mutations, we asked whether a known unlinked suppressor of an intron mutation could suppress the same intron mutation in the *ACT1-CUP1* gene fusion.

prp 16-1 is a dominant suppressor of the *ACT1* branch point mutation, A259C; *prp 16-1* was isolated in a genetic selection for extragenic suppressors of the splicing defect conferred by the A259C mutation in an *ACT1-HIS4* gene fusion (Couto, et al., 1987). A plasmid containing the *prp16-1* allele was cotransformed with the A259C *ACT1-CUP1* gene fusion into the Δcup1 strain which contain a wild-type chromosomal copy of *PRP 16*. Yeast containing the *prp16-1* allele were able to grow on media containing 0.18mM copper while yeast transformed with a control plasmid grew only on media containing 0.10mM copper (Figure 3). Primer extension analyses also are consistent with suppression; approximately two-fold more mature mRNA is produced in the presence of the second- site suppressor (Figure 4B, Table 2).

5' splice site competition assay: Next, we asked if this new gene fusion system could be used to assess the relative contributions of different nucleotides of the 5' splice site. An extremely powerful strategy for addressing the role of sequences in splicing has been to directly compare the activity of two sites in competition (Reed and Maniatis, 1986). This type of competition assay can reveal contributions to splice site choice of components which affect this decision but are not normally rate-limiting for splicing. For example, a 3' splice site competition assay was essential for demonstrating a role for a modestly conserved polypyrimidine tract in *S.cerevisiae*. (Patterson and Guthrie, 1991).

We thus established a 5' splice site competition assay to examine the role of each of the nine nucleotides of the 5'ss consensus sequence. The 5' splice site of *ACT1-CUP1* was replaced with a set of duplicated 5' splice sites separated by 13 nucleotides (Figure 5A). In the first construct (reading frame I, Figure 5B), cleavage at the downstream 5' splice site places the *ACT1* start codon in frame with the *CUP1* coding sequence. In a parallel construct (reading frame II, Figure 5C), only cleavage at the upstream splice site results in an in-frame fusion mRNA. Thus, by adjusting the reading frame it is possible to monitor use of either 5'ss by growth on copper containing media.

When both sites in our competition construct are wild-type, cleavage at the upstream site is favored. In reading frame I, when cleavage at the downstream site results in an in-frame mRNA, the ratio of the mRNAs generated by cleavage at the upstream site relative to the downstream site is only 1.7:1.0. However, if cleavage

at the upstream site results in an in-frame mRNA, then the relative ratio is 6.6:1.0 (Figure 5D). The apparent use of the upstream and downstream sites depends on which of the two mRNAs generated is in-frame; because mRNAs which contain stop codons near their 5' ends have decreased stability (Lossom and Lacroute, 1979).

5' splice site competition mutagenesis: We created a bank of random mutations in the favored upstream 5' ss by doped oligo mutagenesis (Figure 6, see materials and methods). In order to directly select for the mutations which affected 5'ss choice we assumed that any mutation introduced into the highly conserved upstream 5'ss that decreased or abolished use of that site would consequently increase use of the competing unmutagenized downstream wild-type 5'ss. Therefore, the mutated 5'ss competition bank was cloned into reading frame I so that the level of copper-resistance of strains carrying the mutant bank would reflect cleavage at the unmutagenized downstream site. Thus, we were able to isolate mutants by selecting for an increase in copper-resistance rather than screening for a decrease in copper-resistance. Since the Δcup1 strain containing the unmutagenized reading frame II construct can grow at 0.18mM copper, Δcup1 strains containing the mutagenized bank were screened for growth on media containing 0.25mM copper. The DNA plasmids containing the mutagenized 5'ss competition were isolated from the strains and sequenced to identify the mutations.

As summarized in Table 1, all three possible nucleotide changes were isolated at intron positions 1,2,3,4 and 5. No mutations were

isolated at intron position 6 nor at the three exon nucleotides adjacent to the intron (positions -1,-2,-3). All 15 of the mutants at the first five intron positions can grow on some amount of copper >0.18mM, but they vary in their upper tolerance (Table 3). In order to determine whether the differences in growth on copper media reflected the ability of different mutant 5' splice sites to compete against the wild-type 5'ss, RNA was isolated from each of the mutants and analyzed by primer extension with the exon 2 *CUP1* primer. This primer is able to detect the two mature products generated by cleavage at each of the competing 5' splice sites.

Primer extension analyses of mutants at positions 1,2,3 and 5 of the upstream site revealed no mRNA products which result from usage of that site (data not shown). However; since cleavage at this mutated upstream 5'ss results in the production of an out-of-frame mRNA, the failure to detect such products may be due to the rapid turnover of the out-of-frame *ACT1-CUP1* mRNA. In contrast, all three changes at position 4 are able to compete with the wild-type site: wild-type> T4C > T4G > T4A (Table 3).

To increase the sensitivity of the assay, the reading frame of all 15 point mutants was changed so that cleavage at the upstream mutagenized 5' splice site would result in an in-frame (more stable) fusion mRNA. RNA was isolated from yeast transformed with each of these constructs and then analyzed by primer extensions with the exon 2 *CUP1* primer (Figure 7). Indeed, altering the reading frame made it possible to detect low level use of the mutant 5'ss in four cases: intron positions T2C, A3T, A3C and G5C. All three mutations at position 4 are still able to compete and in fact, they compete

even better than before, presumably because cleavage now results in an in-frame stable mRNA.

Mutagenesis of intron position 6 and exon positions -2 and -1: Since no mutants were recovered at intron position 6 or in the three exon positions, we used site-directed mutagenesis to distinguish whether they have no effect in the competition or were simply missed in the original selection.

The last exon 1 position, G-1, was mutated to T (G-1T) and C (G-1C). These changes decrease the potential complementarity of the 5'splice site to the 5' end of U1snRNA (Figure 1). When use of the downstream site is assayed by copper growth, strains containing the mutant plasmids grow at the same copper levels as strains containing plasmids containing the wild-type 5'ss competition (Table 3), thus, these mutants were not detected in our genetic selection. Similarly, neither mutant has an effect when assayed by primer extension (Figure 8).

Exon position -2 was mutated to an A (T-2A). This change increases the potential complementarity of the 5'ss to the regions of both U1 snRNA and U5 snRNA which have been demonstrated to base pair with the 5'ss (Figure 1)(Siliciano and Guthrie, 1988; Seraphin, et al., 1988; Newman and Norman, 1991). Primer extension analyses (Figure 8) and copper-resistance levels of strains containing this mutation (Table 3) demonstrate that the T to A mutation at this position actually increases cleavage at the mutated site relative to the wild-type site. Consequently, it is not

surprising that this change was not detected by the genetic selection scheme.

All three changes were made at intron position 6. When the wild-type T at position 6 is mutated to G or C, primer extension analyses reveal that the mutated 5'ss can still compete with the wild-type 5'ss, although not as well as the wild-type site (Figure 8). Interestingly, these mutated 5' splice sites compete better than changes at positions 1,2,3 and 5. When an A is introduced at this position (T6A), cleavage at the downstream wild-type site apparently results in the production of an unstable message due to the introduction of a stop codon (Lossom and Lacroute, 1979). When use of the downstream wild-type site is assayed by growth on media containing copper, the strains transformed with T6C and T6G can grow at 0.5mM copper while yeast transformed with 6A grow at 0.013mM copper. It is surprising that mutants 6C and 6G were not detected in the genetic selection.

DISCUSSION

Is the *ACT1-CUP1* gene fusion a good reporter to monitor pre-mRNA splicing?: The intron-containing fragment of the yeast *ACT1* gene was fused to *CUP1* and the gene fusion was placed on a 2μ vector under the control of the strong constitutive glyceraldehyde-3-phosphate dehydrogenase promoter (Schena, et al., 1990). A careful analysis of the phenotypes conferred by a number of previously characterized *ACT1* intron mutants illustrates some of

the advantages of the *ACT1-CUP1* system. First, it is extremely sensitive; that is, very low levels of the fusion protein can be detected by copper growth assays. This is in striking contrast to *ACT1-HIS4* fusions. For example, when the branch point mutation, A259C, is introduced into the *ACT1-HIS4* fusion gene, one cannot detect the production of mature mRNA by *in vivo* growth assays on media containing histidinol, although primer extension analyses demonstrate that mature mRNA is produced (Vijayraghavan, et al., 1986). However, a Δcup1 strain containing an *ACT1-CUP1*-A259C fusion can grow on media that contains 0.18mM copper, a concentration at least ten-fold greater than that at which the Δcup1 strain can grow. A second advantage to the *ACT1-CUP1* system, is the large range over which copper growth phenotypes can be monitored. For example, whereas the Δcup1 strain normally grows at 0.013mM copper, the Δcup1 strain containing the wild-type *ACT1-CUP1* fusion grows at 1.0mM copper. Therefore, the system provides about a 100 fold range in which the phenotypes conferred by different intron mutations can be compared. Third, the copper growth phenotypes of different intron mutants generally correlate with the amount of mRNA produced.

Finally, the *CUP1* gene fusion provides a new and powerful tool which can be used to design genetic selections and screens to identify trans-acting components. The sensitivity of the system allows the detection of even small changes in the copper growth phenotypes. It is possible either to directly select for changes which increase the ability of a strain to grow on copper-containing media or to screen for changes which decrease the ability of a strain

to grow on copper-containing media. As just discussed, *CUP1* gene fusions are preferable to *HIS4* gene fusions because of their high sensitivity and large range of detection. In addition, *CUP1* gene fusions have an obvious advantage over *LacZ* gene fusions in that it is possible to directly select for improved growth rather than to screen for increases in ß-galactosidase activity.

Analysis of the role of the nucleotides of the 5'ss in yeast introns: An application of *ACT1-CUP1*.

A 5'ss competition assay was developed in yeast in order to further characterize the roles of the conserved nucleotides of the splice site consensus sequence. A second copy of the 5'ss was introduced 13 nucleotides upstream of the wild-type 5'ss of the *ACT1* gene. This cis-competition construct differs from that of (Goguel, et al., 1991) where the 5' splice sites are separated by 75 nucleotides. In the latter case, cleavage occurs equally at each of the two 5' splice sites. In our competition assay, however, although both of the 5' splice sites are utilized, cleavage at the upstream site is significantly favored over the downstream site, 1.7 to 6.6 fold, depending on the relative reading frames of the two mRNAs. One explanation for this might be that in our construct the upstream site is adjacent to an authentic exon sequence while the downstream site is adjacent to a synthetic exon sequence; in the case of (Goguel, et al., 1991) both sites are flanked by authentic exon sequence. Splice site context has previously been demonstrated to affect splice site choice in mammalian 5'ss competition assays (Reed and Maniatis, 1986; (Nelson, et al., 1990).

Mutagenesis of the upstream 5' splice site: Single point mutations were introduced into each of the conserved nucleotides of the favored upstream 5'ss intron and exon sequences (ATG/GTATGT). A copper growth selection was then used to identify those mutations which reduce cleavage at the upstream site and thereby increase usage of the wild-type downstream site. In this selection, all three point mutation changes were recovered at intron positions 1,2,3,4 and 5. Most changes at intron positions 1, 2 ,3 and 5 resulted in complete loss of cleavage at the mutated site. These are the most highly conserved nucleotides of the consensus sequence (Woolford, 1989) and are complementary to the 5' end of U1snRNA. A base pairing interaction has been demonstrated between yeast U1snRNA and intron positions 1 and 5 (Siliciano and Guthrie, 1988; Seraphin, et al., 1988). The most likely explanation for the effects of the mutations at intron positions 1,2,3 and 5 is thus that these mutations decrease the ability of the 5'ss to stably bind to U1snRNA. Consequently, when these splice sites are competing with a wild-type splice site, U1snRNA binds more stably to the wild-type 5'ss resulting in the commitment of this 5'ss to the splicing pathway. This model is consistent with that derived from the results of mammalian 5'ss competition assays where 5'ss choice appears to be dependent on the overall complementarity of the splice site sequence to U1snRNA (Nelson and Green, 1988; Lear, et al., 1991).

Sub-optimal 5' splice site sequences: Role in regulation of gene expression: A striking correlation is that the three changes, at positions 2 and 3, which retain the ability to compete with the

wild-type splice site, albeit at low levels, are all found in wild-type yeast introns. One gene, *COX5B* (Hodge and Cumsky, 1989) deviates from the 5'ss consensus sequence at position 2 (/GT to /GC) and two genes, *MER2* (/GTAT to /GTTC, (Engebrecht, et al., 1991) and *RPL32* (/GTAT to /GTCA, (Dabeva, et al., 1986)), deviate at position 3. Interestingly, the two genes which contain the position 3 variants also deviate from the consensus sequence at position 4 and these genes are the only two yeast genes whose expression is known to be regulated at the level of splicing. These results thus make it clear that yeast, like mammals, utilize sub-optimal splice site sequences to regulate gene expression.

Intron position 6 influences 5'splice site choice: Intron position 6 is not as highly conserved as intron positions 1,2,3 and 5 (Woolford, 1989) and single point mutations at this position do not confer a detectable phenotype in an otherwise wild-type intron (Seraphin and Rosbash, 1989b). However, a base pairing interaction can be demonstrated between position 6 and U1 snRNA in the presence of a second intron mutation in the branch point consensus sequence (Seraphin and Rosbash, 1989b). Mutations at position 6 in the competition assay compete far better than mutations at positions 1,2 and 5, though not as well as the wild-type 5'ss. This result is consistent with the model that the longer a U1snRNP stays bound to a 5'ss, the greater is the chance that this 5'ss will be utilized. By this view, base pairing between position 6 and the U1snRNA is normally not rate-limiting; however, when competing with a wild-type 5'ss which can base pair with U1snRNA at an

additional position, a role for this nucleotide is uncovered. These results provide strong confirmation of the power of cis-competition assays to reveal the role of sequences not normally rate-limiting for splicing.

A novel role for intron position 4: The 5' end of the *S.cerevisiae* U1 snRNA and the mammalian U1 snRNA are identical, yet the 5'ss consensus sequences differ at position 4. In yeast this position is almost invariably unpaired. If the main criterion for 5'ss recognition in yeast is the ability to maximally base pair with U1snRNA as in mammals (Nelson and Green, 1988; Lear, et al., 1991), then a 5'ss with a position 4 T to A mutation would be expected to be favored over a wild-type site. Our results demonstrate that this is definitely not the case. All three point mutations at this position are able to compete with the wild-type site. In fact, the favored nucleotide at this position is the wild-type T and the mutation which competes the best is a C, another pyrimidine. Therefore, it is likely that intron position 4 is examined independently of the base pairing interaction between the 5'ss and the U1snRNP. Alternatively, the presumed mismatch at position 4, generated when the 5'ss and U1snRNP are base paired, may be important for reasons of helical geometry. In either case, our results indicate that, unlike mammals, the prime determinant of 5'ss choice in yeast is not simply maximal base pairing with U1snRNA.

The role of exon nucleotides in 5' splice site choice: No mutations at any of the three exon positions tested were recovered in the selection. In yeast, the exon nucleotide at position -1 is a G in 53% of the 53 sequenced intron-containing genes (Rymond and Rosbash, 1992). Our analyses of two mutations synthesized by site-directed mutagenesis of position -1 of the upstream 5'ss confirmed that mutations at this position confer no phenotype in the competition assay. In mammals, however, position -1 is very important since this G, along with the first invariant G at intron position 1, is thought to define the cleavage site through its base pairing interaction with the U1snRNA (Aebi, et al., 1987). Our genetic result, together with the sequence information, leads to the conclusion that in yeast, unlike in mammals, complementarity to U1snRNA at position -1 is not very important. Taken together, the intron position 4 and exon position -1 results indicate that, yeast, in addition to complementarity to U1snRNA, yeast must employ some novel means to determine 5' splice site choice.

Exon position -2 is the least conserved nucleotide of the consensus sequence. It is often an A (36%) or a T (26%) (Rymond and Rosbash, 1992). In our competition construct this nucleotide is a T. When it is mutated to A, cleavage at the mutant 5' splice site is increased when in competition with the wild-type splice site. An A at this position is complementary to U1snRNA (see Figure 1). This result is in contrast with our data for position -1, since changes of this nucleotide which should increase complementarity to U1snRNA do not affect 5'ss choice. However, the advantage of an A at position -2 may not be due to a base pairing role with U1 snRNA. Recent

results demonstrate that exon positions -2 and -3 can base pair with U5 snRNA and that this interaction can be important in cleavage site choice (Newman and Norman, 1991). The T to A mutation provides potential complementarity of this 5'ss nucleotide to both U1 snRNA and U5 snRNA. Conceivably, the different effects of mutations at positions -1 and -2 reflect the importance of the complementarity between -2 and U5 snRNA in determining the cleavage site choice. An alternative explanation is that when complementarity is increased between U1snRNA and position -2, the length of the continuous region of potential base-pairing between U1snRNA and the 5'ss increases from 4 to 6 base pairs. In the case of position -1 mutations, the wild-type splice site can base pair at five adjacent positions while the mutated splice sites can only base pair at three adjacent positions.

Summary: The combined power of the 5'ss cis-competition assay and the sensitive *ACT1-CUP1* reporter has allowed the genetic selection of a large set of 5'ss mutants. The competition strategy was essential in uncovering a role for intron positions 4 and 6 and exon position -2, since mutations at these positions have no detectable phenotypes on their own. These findings argue that the *CUP1* fusions will continue to be an extremely useful tool in our genetic analysis of splicing. Moreover, in principle, *CUP1* can be fused to any gene or regulatory region in order to study other basic cellular processes. Presently, *CUP1* gene fusions are being used to study mRNA stability (R. Parker, personal communication), transcription (J. Lefstein and K. Yamamoto, personal communication)

and protein translocation (D. Ng and P. Walter, personal communication).

Figure 1. Model of the proposed base-pairing interactions between the 5' splice site and U1 and U5 snRNPs. The numbers beneath the 5'ss designate each of the positions of the consensus sequence.

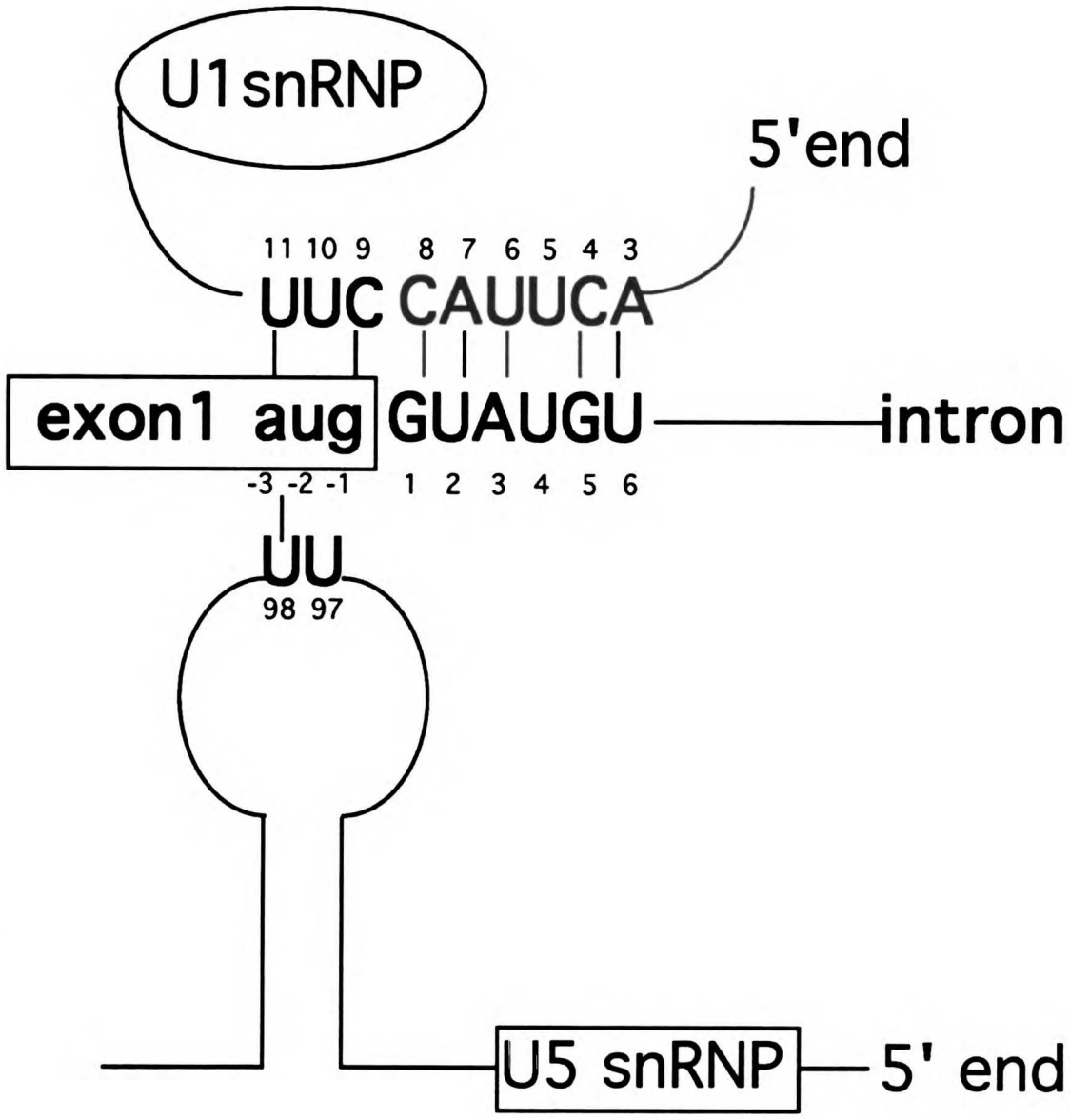

U1 snRNP
5'end
11 10 9 8 7 6 5 4 3
UUC CAUUCA
exon1 aug
GUAUGU
intron
-3 -2 -1 1 2 3 4 5 6
UU
98 97
U5 snRNP
5' end

Figure 2. Map of the *ACT1-CUP1* gene fusion. B, S, X, K, A and SI refer to sites for restriction endonucleases *Bam*H I, *Sac* I, *Xho* I, *Kpn* I, *Asp* 718, and *Sal* I, respectively.

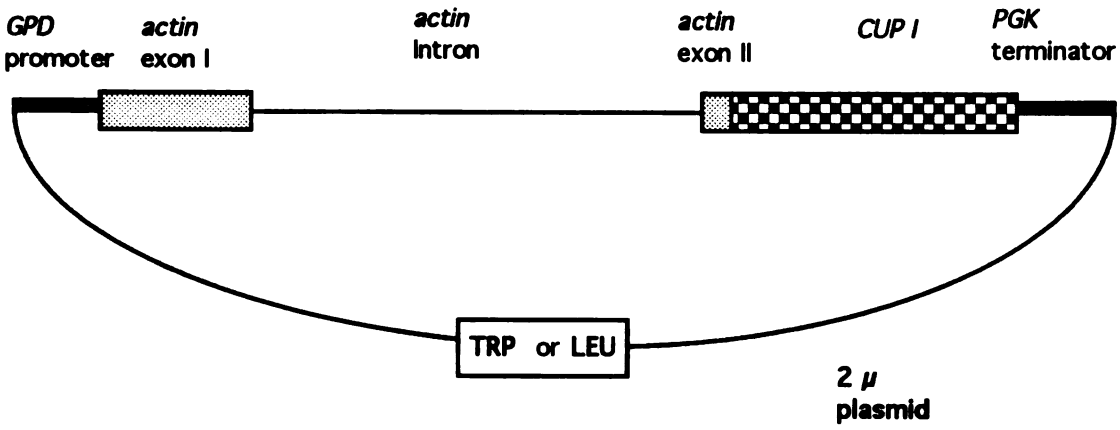

GPD
promoter
actin
exon I
actin
Intron
actin
exon II
CUP I
PGK
terminator
TRP or LEU
2 μ
plasmid

Figure 3. Copper growth phenotypes of strains containing intron mutations. Strains containing each of the point mutations were assayed for copper-resistance levels by the liquid spotting method (see materials and methods). The specific mutation carried by each strain is designated in the large circle. The specific copper concentrations of the plates are designated in the smaller circles.

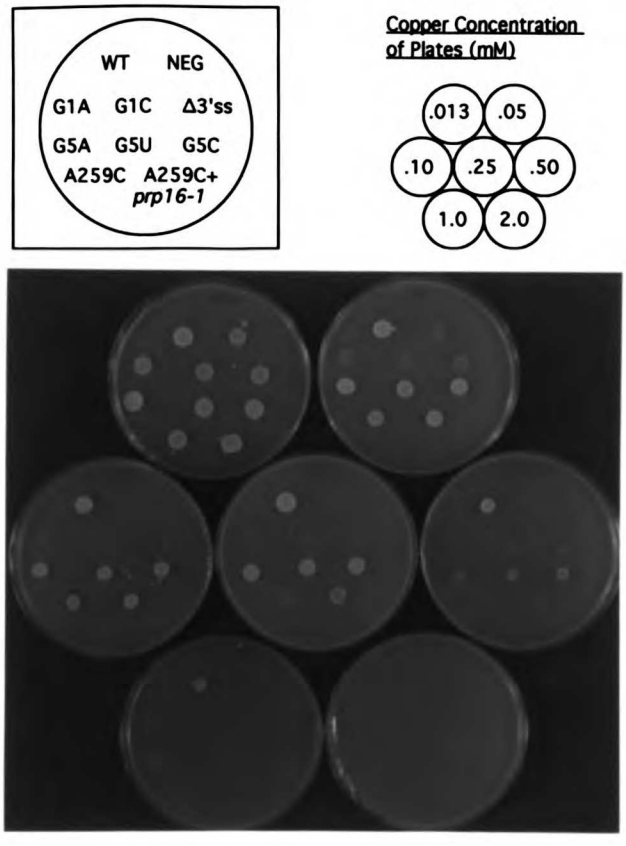
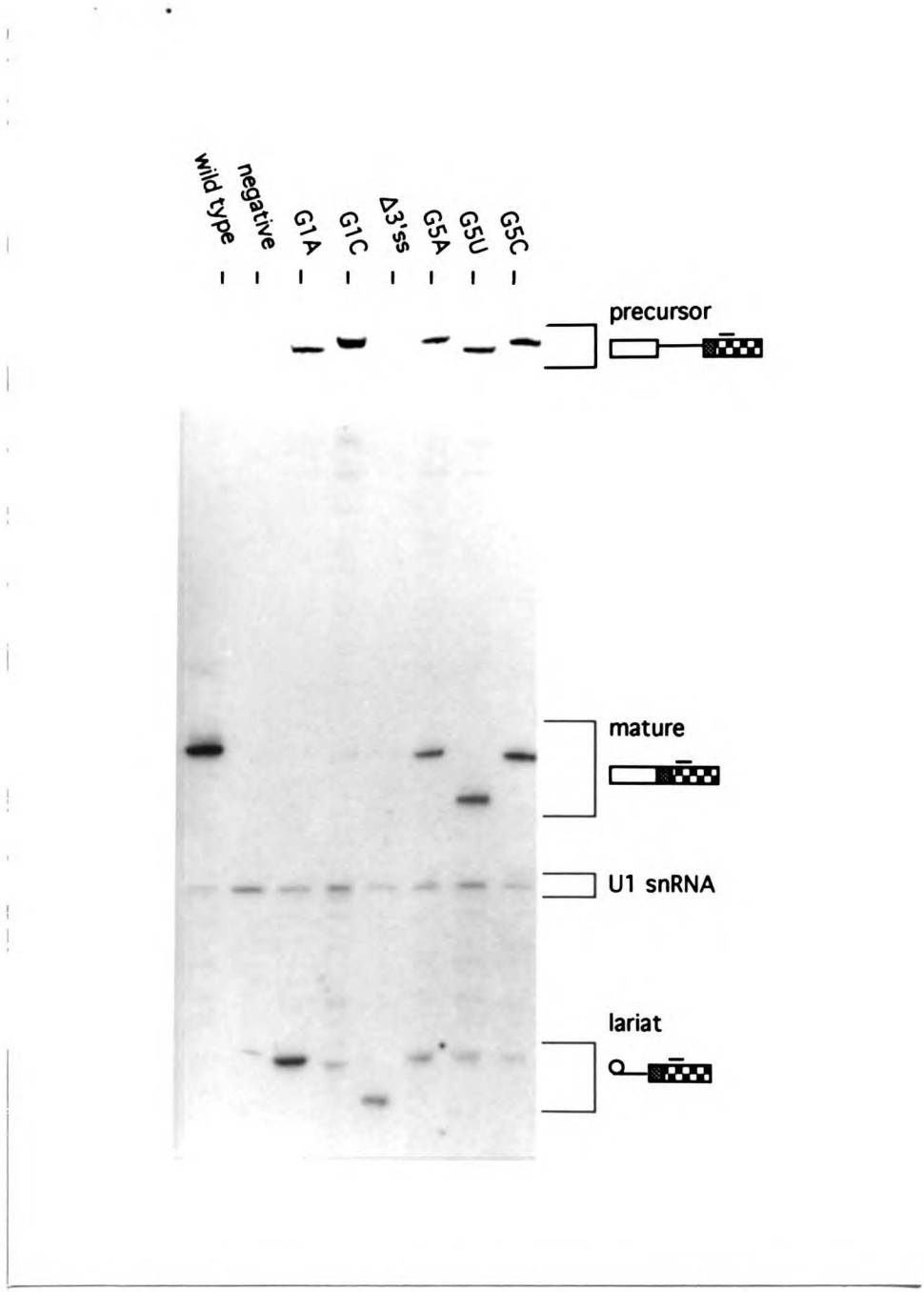

WT
NEG
G1A
G1C
Δ3'ss
G5A
G5U
G5C
A259C
A259C+
prp16-1
Copper Concentration of Plates (mM)
.013
.05
.10
.25
.50
1.0
2.0

Figure 4. Primer extension analysis of intron mutations. **(A)** Primer extension analysis of 5'ss and 3'ss mutations. RNA was prepared from strains carrying the intron mutation indicated at the top of each lane. Wild type indicates a strain transformed with an *ACT1-CUP1* construct containing the unmutagenized intron while negative indicates a strain that does not carry the *ACT1-CUP1* gene fusion. The sizes of the precursor and mature bands vary due to differences in the lengths of the 5' untranslated regions resulting from variations in the cloning strategies for moving the intron mutations (see materials and methods). **(B)** Primer extension analysis of suppression of a branch point mutation. RNA was isolated from a strain carrying the A259C intron mutation and either prp16-1 (pM3) or a control plasmid (pSE358). The possible products of the primer extension reactions are diagrammed on the side of the gel. Each lane was primer extended with *CUP1* and the U1 snRNA primers. The location of the exon 2 primer (3'Cup) is noted on the diagrams by a line above exon 2.


wild type
negative
G1A
G1C
Δ3'ss
G5A
G5U
G5C
precursor
mature
U1 snRNA
lariat

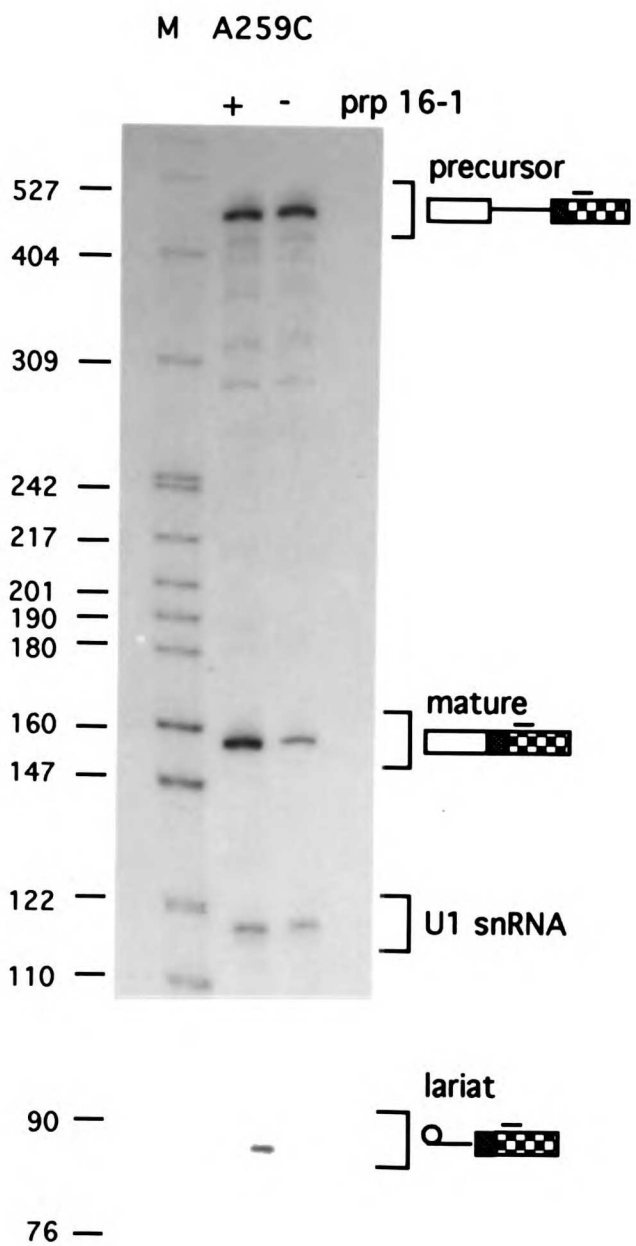
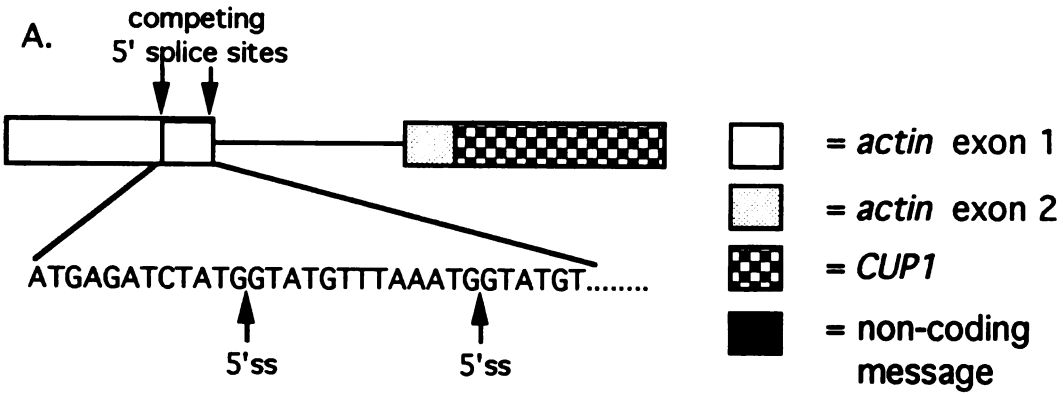
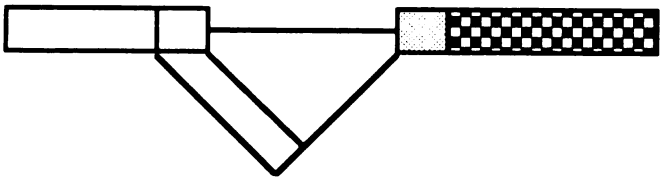

M A259C
+ - prp 16-1
527
404
309
242
217
201
190
180
160
147
122
110
90
76
precursor
mature
U1 snRNA
lariat

Figure 5. 5' splice site competition assay. **(A)** General map of the 5'ss competition assay. The exact sequence of the 5'ss duplication is shown. The first ATG of the sequence is the new start codon of the fusion gene followed by the *Bgl* II restriction site and the first of the 5' splice sites. The second 5'ss is located 13 nucleotides downstream from the first.

(B) and **(C)** Schematic representations of the mature messages generated as a result of cleavage at each of the two possible 5' splice sites. The relative reading frame of the competition constructs determines which of the two possible mature mRNAs will result in the production of a functional copper binding fusion protein. (D) Primer extension analysis of the reading frame I and reading frame II 5'ss competition constructs.



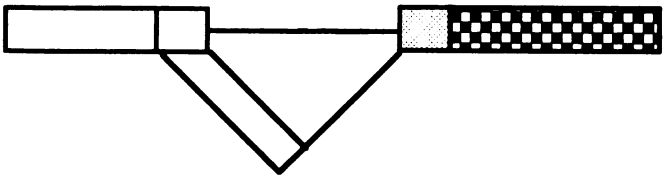
B. Reading Frame I



In-Frame = Copper-Resistant

Out-of-Frame = Copper-Sensitive

C. Reading Frame II (filled in *Asp* 718)



Out-of-Frame = Copper-Sensitive

In-Frame = Copper-Resistant

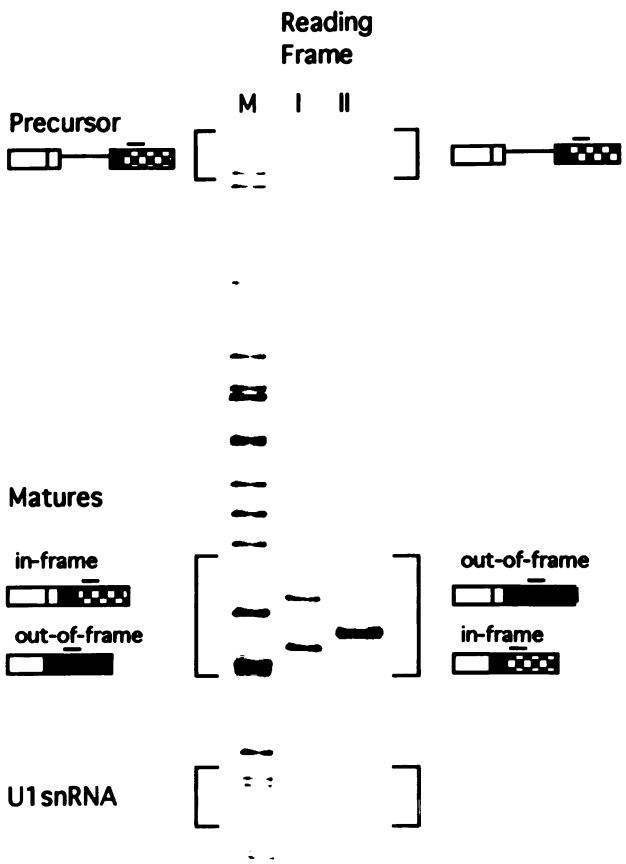

Reading
Frame
M
I
II
Precursor
Matures
in-frame
out-of-frame
out-of-frame
in-frame
U1snRNA

Figure 6. Schematic representation of the strategy used to select for 5'ss mutations which affect the competition assay. (see materials and methods for details.)

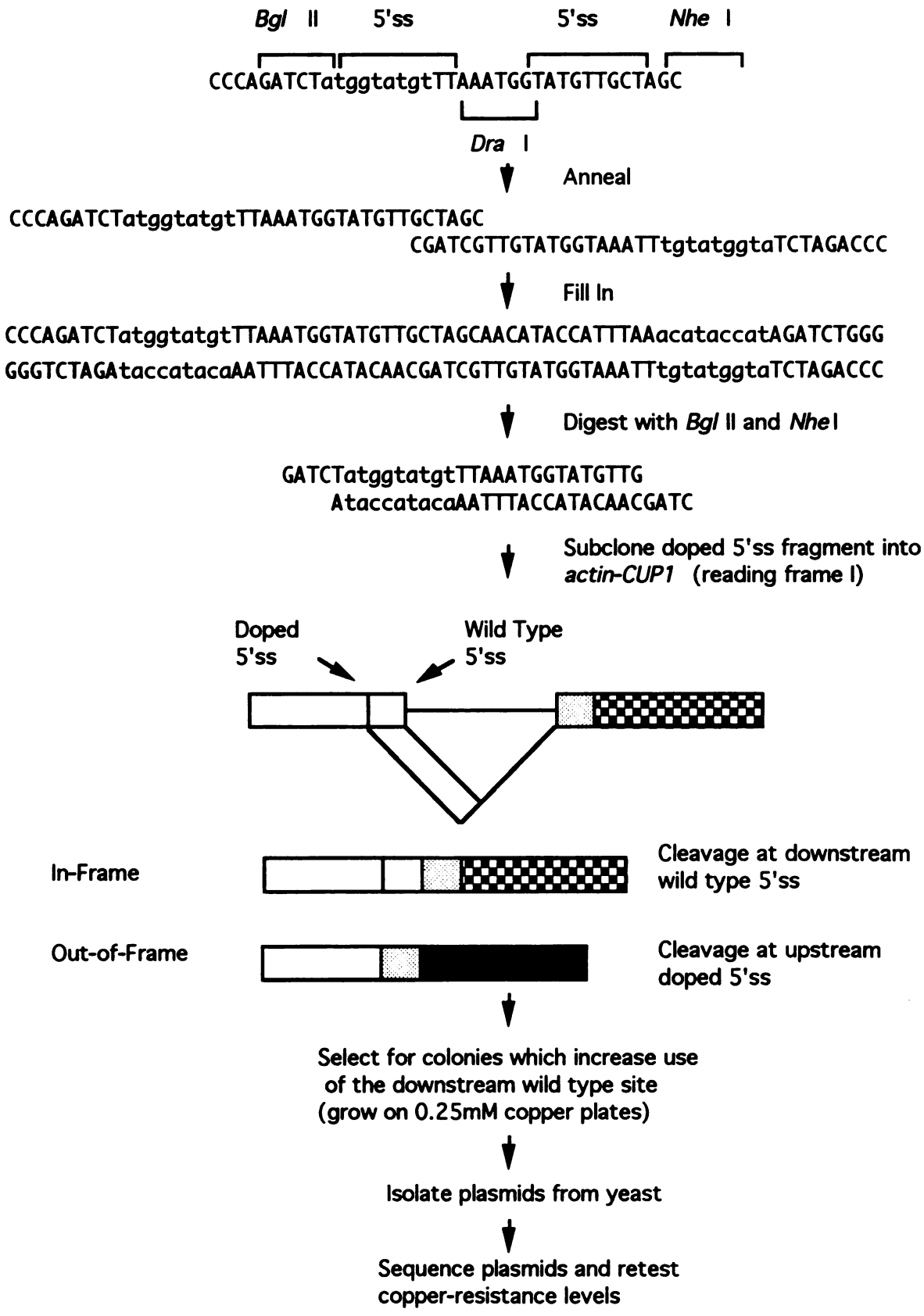

Bgl II
5'ss
5'ss
Nhe I
CCCAGATCTatggtatgtTTAAATGGTATGTTGCTAGC
Dra I
Anneal
CCCAGATCTatggtatgtTTAAATGGTATGTTGCTAGC
CGATCGTTGTATGGTAAATTtgtatggtaTCTAGACCC
Fill In
CCCAGATCTatggtatgtTTAAATGGTATGTTGCTAGCAACATACCATTTAAacataccatAGATCTGGG
GGGTCTAGAtaccatacaAATTTACCATACAACGATCGTTGTATGGTAAATTtgtatggtaTCTAGACCC
Digest with Bgl II and Nhe I
GATCTatggtatgtTTAAATGGTATGTTG
AtaccatacaAATTTACCATACAACGATC
Subclone doped 5'ss fragment into
actin-CUP1 (reading frame I)
Doped
5'ss
Wild Type
5'ss
In-Frame
Cleavage at downstream
wild type 5'ss
Out-of-Frame
Cleavage at upstream
doped 5'ss
Select for colonies which increase use
of the downstream wild type site
(grow on 0.25mM copper plates)
Isolate plasmids from yeast
Sequence plasmids and retest
copper-resistance levels

Figure 7. Primer extension analysis of 5'ss competition mutations. RNA was prepared from strains carrying the 5'ss competition mutation indicated at the top of each lane. The 5'ss competition assays shown here are all in reading frame II. The possible products of the primer extension reactions are diagrammed on the side of the gel. Each lane was primer extended with the *CUP1* and the U1 snRNA primers. The location of the exon 2 primer (3'Cup) is noted on the diagrams.

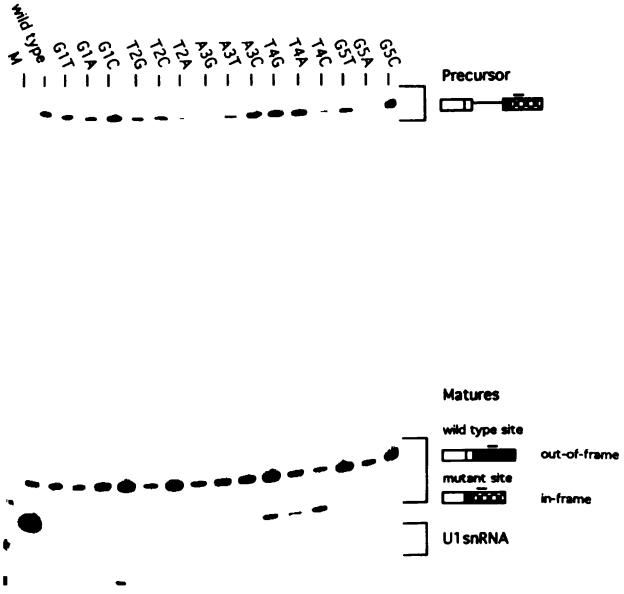

wild type
M
G1T
G1A
G1C
T2G
T2C
T2A
A3G
A3T
A3C
T4G
T4A
T4C
G5T
G5A
G5C
Precursor
Matures
wild type site
out-of-frame
mutant site
in-frame
U1snRNA

Figure 8. Primer extension analysis of positions -2, -1 and 6 mutations. RNA was prepared from strains containing the 5'ss competition mutation indicated at the top of each lane. The 5'ss competition assays shown here are all in reading frame I. The products of the primer extension reactions are diagrammed on the side. Each lane was primer extended with the *CUP1* and the U1 snRNA primers. The location of the exon 2 primer (3' Cup) is noted on the diagrams.

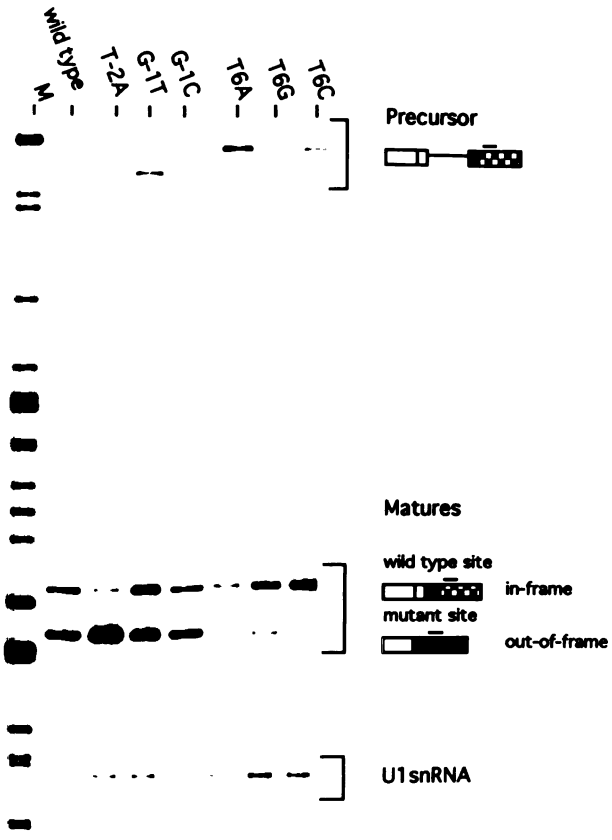

M
wild type
T-2A
G-1T
G-1C
T6A
T6G
T6C
Precursor
Matures
wild type site
in-frame
mutant site
out-of-frame
U1snRNA

TABLE 1

Summary of the number of specific mutations recovered at each position of the degenerate 5' splice site

5' splice site	Mutation			
position	G	A	T	C
-3	0	WT	0	0
-2	0	0	WT	0
-1	WT	0	0	0
1	WT	1	3	1
2	1	1	WT	1
3	1	WT	6	1
4	1	1	WT	3
5	WT	1	4	3
6	0	0	WT	0

TABLE 2

Summary of copper-growth and primer extension phenotypes of intron point mutations

Intron Mutation	Copper Growth	Mature/U1SnRNA[a]
wild type	1.0mM	32.5
no plasmid	0.013mM	0.0
G1A	0.013mM	0.0
G1C	0.013mM	0.0
Δ3'ss	0.013mM	0.0
G5A	0.30mM	2.4
G5U	0.30mM	3.4
G5C	0.50mM	6.9
A259C	0.15mM	1.6
A259C+ prp16-1	0.25mM	2.7

[a] the ratio in this column indicates the amount of mature mRNA relative to the amount of U1 snRNA.

TABLE 3

Summary of copper-growth and primer extension phenotypes of 5'ss competition mutations

INTRON MUTATION	COPPER GROWTH[a]	WILD TYPE/ MUTANT[b]
wild type	0.18mM[c]	1.0/1.7[c]
T-2A	0.05mM	1.0/9.7
G-1T	0.18mM	1.0/1.6
G-1C	0.18mM	1.0/1.9
G1T	0.50mM	
G1A	0.50mM	
G1C	0.50mM	
T2G	0.50mM	
T2C	0.25mM	
T2A	0.50mM	
A3G	0.50mM	
A3T	0.50mM	
A3C	0.50mM	
T4G	0.25mM	1.0/.08
T4A	0.25mM	1.0/.06
T4C	0.50mM	1.0/.25
G5T	0.50mM	
G5A	0.50mM	
G5C	1.0 mM	
T6A	0.013mM	
T6G	0.50mM	1.0/.26
T6C	0.50mM	1.0/.11

[a] copper growth reflects the actin-CUP1 fusion protein produced as a result of cleavage at the downstream unmutagenized 5'ss.

[b] the ratios in this column indicate the amount of mature message generated as a result of cleavage at the unmutagenized downstream 5'ss relative to the mutagenized upstream 5'ss. Cleavage at the unmutagenized 5'ss results in an in-frame mRNA while cleavage at the mutant 5'ss results in an out-of-frame mRNA.

[c] both of the 5' splice sites in this competition are wild type.

CHAPTER 3

Effects of U1 snRNA point mutations on 5' splice site mutations

U1 snRNA was the first factor demonstrated to interact with the 5' splice site. As discussed in Chapter 1, these two RNAs base pair with one another and this interaction is important for the first step of splicing. 5' splice site mutations can be suppressed by compensating mutations in U1 snRNA, however, the restoration of complementarity only partially suppresses the splicing defect. Thus, other factors are involved in either establishing or stabilizing the U1 snRNA·5' splice site interaction.

In order to further analyze the mechanism by which compensating U1 snRNA mutations suppress the decrease in splicing efficiency conferred by intron mutations, I decided to examine the ability of U1 compensating mutations to suppress intron mutations in two new situations: (1) when the mutated pre-mRNA is overexpressed and (2) when a mutated 5' splice site is competing with a wild-type 5' splice site. In addition, I characterized the phenotypes conferred by several U1 snRNA alleles that are mutated in the region that is complementary to the 5' splice site. Interestingly, these studies, like previous studies reviewed in Chapter 1, indicate that other factors are involved in the process of 5' splice site recognition. Furthermore, these studies suggest that U1 snRNA is not the first factor to interact with the 5' splice site.

Suppression of 5' splice site mutations by compensating U1 mutations. Using low copy (cen) ACT1-HIS4 and ACT1-LacZ reporter plasmids, Siliciano and Guthrie (1988) had demonstrated that restoration of complementarity between U1 and the 5' splice site increases the efficiency of the first catalytic step.

Suppression is semi-dominant; that is, suppression is observed in a strain that contains both a wild-type and mutant copy of U1, but suppression increases in a strain homozygous for the U1 mutation (Siliciano and Guthrie, 1988; Seraphin, et al., 1988).

As shown in Chapter 2, the phenotypes conferred by intron mutations in the high copy ACT1-CUP1 reporter system were similar to those phenotypes observed when the same mutations were studied on low copy reporters. Thus, I expected that compensating U1 mutants would be able to suppress the decrease in splicing efficiency conferred by 5' splice site mutations when the 5' splice site mutations were expressed in the high copy ACT1-CUP1 reporter system. Surprisingly, however, when the intron position 1 or 5 mutations were cloned into the high copy (2μ) ACT1-CUP1 reporter system, suppression by compensating U1 mutations was not observed if a wild-type copy of U1 was also expressed in the cells. Suppression by the compensating U1 mutants was only observed if the strain was homozygous for the compensating U1 suppressor. Notably, even in these cases, suppression could not be detected by copper resistance assays; but was observed only by primer extension analyses (Table 1). The failure to observe suppression by copper resistance assays may be due to the poor growth of U1-C4U and U1-C8U suppressor strains, resulting in a decrease in the apparent levels of copper resistance.

The inability of the U1 mutations (in a heterozygous background) to suppress the 5' splice site changes could be the result of expressing the new reporter genes at very high levels, i.e. on 2μ plasmids under the control of the very strong constitutive GPD

promoter. Alternatively, the lack of suppression could be due to the design of the new gene fusion, since almost all of the actin exon 2 sequence had been deleted from the new gene fusion. To distinguish between these possibilities, the intron G1A and G5A ACT1-CUP1 mutant reporters were subcloned onto low copy (cen) plasmids. Under these conditions the U1 suppressors functioned in the presence or absence of a wild type copy of U1 snRNA, as assayed by primer extension and copper resistance (Table 1). Thus, the expression level of the reporter gene influences the suppression by U1.

The variable in these experiments was the amount of mutant pre-mRNA; the absolute amount of suppressor and wild-type U1 probably did not change. One explanation is that some limiting factor, factor X, has been titrated by excess pre-mRNA. According to this model, factor X (1) binds preferentially to wild type 5' splice site sequences and (2) helps mutant U1 snRNAs recognize or bind to a 5' splice site. When the mutant 5' splice site pre-mRNA is expressed at high levels and factor X is limiting, the probability that factor X will bind to the mutant 5' splice sites would be low. Thus, compensating changes in U1 would be unable to suppress intron mutations because the mutant U1 could not recognize or bind to the pre-mRNA. In a strain homozygous for the suppressor U1, the mutant U1 does not have to compete with the wild-type U1 for binding to the 5' splice site and consequently is no longer dependent on the limiting factor X.

Is there a factor X involved in 5' splice site recognition? A family of splicing proteins, known as SR proteins for characteristic dipeptides, are now known to be involved in 5' splice site selection in higher eucaryotes (Krainer et al., 1990; Zahler et al., 1993). Interestingly, several members of this family have been shown to commit pre-mRNA to the splicing pathway in the absence of U1 (Fu, 1993). Recent evidence has led to the proposal that these proteins help U1 snRNA anneal to the 5' splice site (I. Eperon, unpublished data). In addition, ASF/SF-2, the first identified SR protein, has been shown to bind to the 5' splice site sequence in a sequence specific manner. Point mutations in the 5' splice site consensus sequence can reduce binding of ASF/SF-2 to the splice site by as much as a factor of 100, with the largest effects observed in competition assays (Zuo and Manley, 1993). No member of this family has yet been identified in *S. cerevisiae*; however, the above observations suggest that a protein with similar functional properties might exist.

A genetic screen, such as outlined in Figure 1, could be used to identify such a factor. The premise of this screen is that overexpression of the limiting factor X should result in increased cleavage at a mutant 5' splice site (e.g., G5A) when the mutant pre-mRNA is expressed at high levels in the presence of both the compensating suppressor U1 (U1-C4U) and wild-type U1.

Screens for high copy suppressors have relied on expressing the library on a high copy number (2μ) plasmid or under the control of the strong inducible GAL promoter. Since the mutant ACT1-CUP1 is being expressed on a high copy (2μ) plasmid, it would be best to

introduce the yeast library on a low copy (cen) plasmid rather than introduce a second high copy (2μ) plasmid into the cell. That is, a cell containing two 2μ plasmids amplifies the plasmids to the copy number at which the first 2μ plasmid was maintained and will randomly choose the copy numbers of each of the two plasmids (Broach and Volkert, 1991). Thus, the cen-GAL driven library should be introduced into a wild type strain which contains the G5A ACT1-CUP1 reporter on a 2μ plasmid in addition to the compensating U1 mutant (U1-C4U). The transformants should be screened for increased copper resistance that is dependent on both the presence of the G5A ACT1-CUP1 reporter and U1-C4U.

In addition to searching yeast libraries for factor X, this screen can also be used to test if known proteins can confer the desired phenotype. For example, the Drosophila SRp55 protein under the control of the GAL promoter has been overexpressed in yeast (M. Roth, personal communication). Thus, it is possible to now test if this Drosophila protein can affect 5' splice site selection in yeast.

Phenotypes conferred by viable U1 point mutations. The experiments described above led to the conclusion that it is only possible to monitor suppression of 5' splice site mutations on high copy reporters by compensating changes in U1 if the U1 mutants are viable when they are the sole U1 species in the cell. The U1-C4U and U1-C8U alleles are viable, while all other changes at U1 positions 4 and 8 are inviable (Siliciano and Guthrie, 1988; Seraphin, et al., 1988) as is the U1-A6U mutant (K. Nandabalan, personal communication). Presumably, the viability of the C to U mutations

in U1 is because these positions of U1 can form U·G base pairs with the 5' splice site. However, the U1-C4U and U1-C8U mutations confer growth defects presumably due to a decrease in the strength of the U1·5' splice site base-pairing interaction.

Unlike U1 positions 4 and 8, this position is not complementary to the corresponding nucleotide of the 5' splice site. When U1 was mutated at position 5 to increase complementarity to the 5' splice site (U1-U5A), strains homozygous for this mutation are viable; however they grow at a decreased rate, at all temperatures tested, relative to a wild-type U1 strain. Two additional U1 alleles, U1-U5G and U1-U5C, were generated. Like the U1-C4U, U1-U5A and U1-C8U alleles, these alleles are viable. Table 2 summarizes the growth phenotypes conferred by each of these mutations at different temperatures.

The viable U1 mutants appear to fall into two classes. Class 1 is comprised of U1-C4U and U1-C8U which are each cold and temperature-sensitive. Class 2 is composed of the three point mutations at U1 position 5, which do not display conditional growth phenotypes. U1-U5C grows like the wild type U1 strain, while the U1-U5A and U1-U5G strains are slightly sick at all temperatures. Interestingly, the U1-C4U and U1-U5G strains grow very similarly at 30°C, but the relative growth rate of U1-C4U decreases at both higher and lower temperatures (Table 2). Strikingly, class 1 is composed of mutants at U1 positions 4 and 8, positions that normally base pair with the 5' splice site, while class 2 is composed of mutants at U1 position 5, which does not normally base pair with U1.

Perhaps the decrease in the potential base-pairing interaction between either U1-C4U or U1-C8U and the 5' splice site results in the conditional phenotypes of these mutants. Thus, by suppressing the conditional phenotypes of U1-C4U or U1-C8U strains, it might be possible to genetically isolate factors involved in stabilizing the U1·5' splice site interaction, possibly factor X or a U1 snRNP protein that interacts with factor X. Conversely, cold-sensitive phenotypes have, in several cases, been shown to be the result of the hyperstabilization of an intermediate (Dammel and Noller, 1993) and recent experiments indicate that the U1·5' splice site interaction must be disrupted or modified before the addition of the other snRNPs (Konforti, et al., 1993). Thus, suppression of the cold-sensitive phenotype conferred by the U1-C4U and U1-C8U strains, might result in the identification of factors involved in modifying the U1/5' splice site interaction so that the other snRNPs can bind to the pre-mRNA. Possible candidates, in this case, are PRP5, a splicing protein which has helicase-like domains, as well as PRP9, PRP11 and PRP 21, proteins all required for U2 addition (Ruby, et al., 1993).

Restoration of complementarity between U1 and competing mutated 5' splice sites does not increase cleavage at mutated sites. Mutational studies of the 5' splice site have demonstrated that this sequence is involved in several steps of the splicing pathway including: (1) commitment to the splicing pathway by base-pairing with U1 snRNA (2) 5' cleavage site choice and (3) determination of 3' splice site cleavage by inspection of the

branched intermediate (for review see Chapter 1). The 5' splice site competition assay described in Chapter 2 was designed to assess the relative contributions of different nucleotides in the 5' splice site consensus sequence. This assay demonstrated that mutations in the intron portion of the splice site sequence caused step 1 defects that decreased or abolished spicing at the mutated splice site. The most likely explanation for these observations was that the mutant 5' splice sites couldn't compete with the nearby wild type site for U1.

A straightforward way to test this hypothesize was to determine whether compensating changes in U1, predicted to restore base-pairing with a mutated site, could increase use of the competing mutated site. As discussed above, strains homozygous for either U1-C4U or U1-C8U partially suppress the decrease in the efficiency of step 1 conferred by intron position 5 or 1 mutations, respectively, by restoring the complementarity of these two RNAs. Thus, the complete set of competing intron position 5 mutations was introduced into the U1-C4U strain and the complete set of competing intron position 1 mutations was introduced into the U1-C8U strain.

Surprisingly, no cleavage was observed at the mutated 5' splice sites when assayed in the presence of a compensating U1 mutation as assayed by either copper resistance or primer extension assays (data not shown), even when all the U1 snRNA expressed in the cell contained the compensating mutation. Thus, the inability of the G5A and G1A sites to compete with the nearby wild-type 5' splice site is not (simply) a result of their decreased complementarity to U1.

U1-C8U strain exhibits a loss of mRNA. Interestingly, primer extension analyses of RNA isolated from the U1-C8U strain carrying the competition reporters revealed a loss of mature ACT1-CUP1 mRNA (Figure 2). That is, the amount of mature ACT1-CUP1 mRNA in the mutant U1-C8U strain was much lower than the amount of mature ACT1-CUP1 mRNA in the wild-type U1 strain. This phenotype was not observed when the strain was transformed with a plasmid which only contains either a single wild-type or mutant 5' splice site. One intriguing interpretation of this data is that U1-C8U has trouble binding to the 5' splice site competition pre-mRNA and thus can't commit this message to the splicing pathway, resulting in the degradation of the pre-mRNA. It is worth noting that even in a U1 wild-type strain, when two 5' splice sites are in close proximity, the efficiency of the first catalytic step of splicing is decreased. Similarly, previous mammalian studies have shown that if two 5' splice sites are in close proximity they can inhibit step 1 in vitro (Nelson and Green, 1988).

Effects of creating complementarity between U1 position 5 and intron position 4 in the competition assay. Strains homozygous for each of the U1 position 5 mutants were transformed with the competition reporters mutated at intron position 4. The primer extension data are shown in Figure 3 and quantitated in Table 3. The values in table 3 indicate the amount of cleavage at the upstream site (U4WT, U4C, U4G or U4A) relative to cleavage at the downstream wild type site. As can be seen by reading across each line of the Table, in each of the strains the order of 5' splice site

choice is always U4WT>U4C>U4G>U4A. Surprisingly, however, the ability of an intron position 4 mutation to compete against a wild type site is increased when complementarity is created between U1 position 5 and intron position 4, as can be determined by reading down each column of the Table. For example, in a U1-U5C strain, the intron U4G mutant can compete better against the wild type site than in any other strain. However, even in the U1-U5C strain, the upstream U4G mutant site cannot compete as well as the upstream U4C mutant site with the wild-type downstream site.

These results are interesting because they demonstrate that under select conditions a base-pairing interaction between U1 position 5 and intron position 4 can result in increase in use of a splice site. This result was not expected for these reasons. (1) In a U1 wild-type strain, in the 5' splice site competition assay, the stronger the potential base-pairing interaction between the U at position 5 of the wild-type U1 and position 4 of the 5' splice site, the less likely this site could compete with the wild type site (wild-type>4C>4G>4A, see Chapter 2). (2) Mutations in U1 at position 5, which increase complementarity to the 5' splice site consensus sequence, confer growth defects on the cell. Thus, it appears that both intron position 4 and U1 position 5 normally play roles in splicing other than the U1/5' splice site base-pairing interaction.

What is the competition assay measuring? Compensating changes are known to suppress the decrease in efficiency conferred by intron position 1 and 5 mutations, however the same

compensating U1 changes cannot suppress the same mutant 5' splice sites when in competition with a nearby wild-type site. These observations suggest that the competition is measuring a step in splicing distinct from commitment of the 5' splice site to splicing by base-pairing with U1. In principle, the competition assay could be monitoring a step prior to (the "recruitment model") or after U1 commitment (the "splice site region vs. cleavage site location" model), or possibly both.

The "recruitment model" proposes that a factor examines the 5' splice site sequence or region prior to U1. This factor searches for the sequence with the best match to the 5' splice site consensus sequence and then directs U1 to the 5' splice site. Thus, even when compensating changes are introduced into U1 the mutant U1 is never directed to the competing mutated 5' splice site. In the simplest case, the "recruitment factor" is the same as the factor X proposed earlier to explain dosage-dependent suppression in non-competition constructs. In this case, over-expressing factor X should allow the mutant 5' splice site to be recognized, bind U1 and consequently compete.

The "splice site region vs. cleavage site location" model proposes that other factors examine the 5' splice site after U1. That is, U1 binds to the 5' splice site and commits a region of the intron that includes a 5' splice site to the splicing pathway. Other factors are then responsible for choosing the cleavage site. In this model, the U1 position 4 and 8 mutations might increase the initial recognition of the competing intron position 1 or 5 mutations, but other factors, perhaps U5 and U6, choose the exact cleavage site.

Indeed, recent experiments demonstrate that mutations in both U5 and U6 can activate cleavage at the mutated splice sites (J. Staley and C.L.).

How can the suppression of intron position 4 mutants by compensating changes in U1 (position 5) be explained by these models? Curiously, as described above, the restoration of complementarity between U1 snRNA and the 5' splice site does not activate cleavage at a mutated 5' splice site; however, the creation of a potential base pair between two normally non-complementary positions can increase cleavage at a mutated 5' splice site. In terms of the "recruitment model," it is note worthy that when complementarity is created, between intron position 4 and U1 position 4, the potential stability of the U1·5' splice site interaction is greater than in the normal situation. Perhaps, this increase in stability allows the suppressor U1 to bind to the mutant 5' splice site in the absence of the proposed "recruitment" factor.

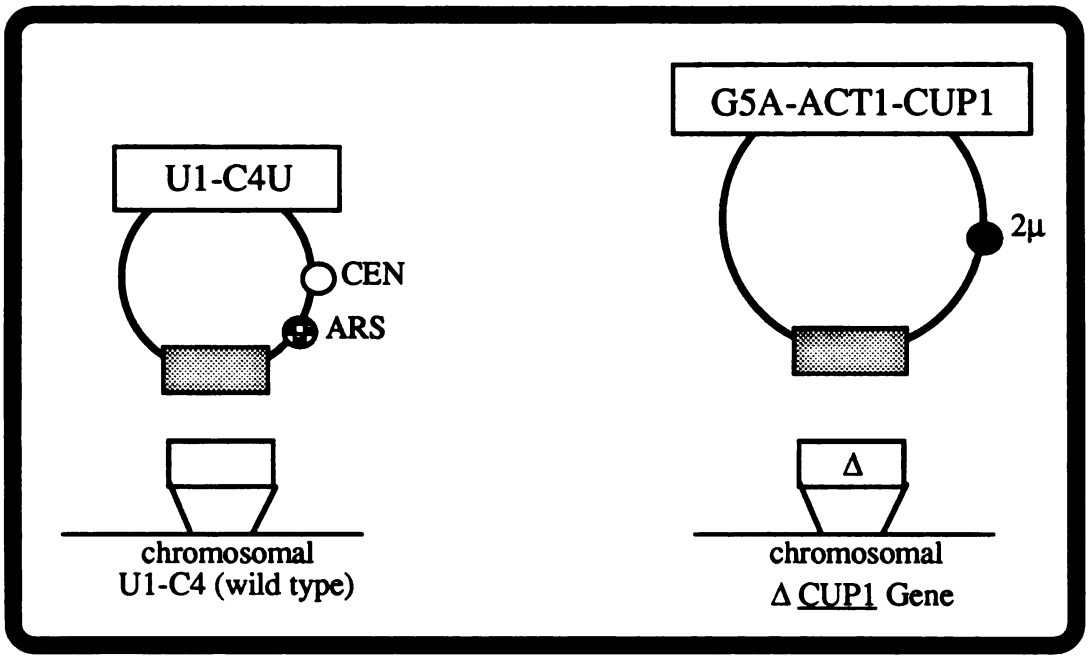
Interestingly, even though the U1 position 5 suppressors can increase cleavage of a mutated 5' splice site containing the complementary intron position 4 mutation, the relative amount of cleavage is always the same (U4WT>U4C>U4G>U4A). That is, even though cleavage is increased at the competing U4G site, in the U1-U5C strain, this site (U4G) still cannot compete as well as the U4C intron mutation against the wild type 5' splice site. Thus, the competition assay must be monitoring the interactions of intron position 4 with at least one other factor besides U1. This other interaction ensures that no matter what changes are made in U1, the

order of use of the competing 5' splice site mutated at position 4 will always be wild-type>U4C>U4G>U4A. Thus, in terms of the "splice site region versus cleavage site location" model, the U1 position 5-intron position 4 might be important in defining the splice site region while the interaction of intron position 4 with the factor other than U1 snRNA is important in choosing the exact cleavage site location.

In summary. The observations presented in this chapter indicate that the process of 5' splice site recognition in *S. cerevisiae* , like mammals, cannot solely be defined by a U1 snRNA·5' splice site base-pairing interaction. Rather, other factors appear to be involved in this process. The characterization of the interactions of U1 snRNA with the overexpressed mutated pre-mRNA and the 5' splice site competition assay make several predictions as to the roles of these factors which can now be used to design genetic strategies to identify such factors.

Figure 1. Genetic scheme to identify factor X. The yeast drawn is deleted at the chromosomal CUP1 locus. The strain carries a plasmid which expresses the G5A ACT1-CUP1 reporter. The strain is a haploid, but is heterozygous for U1 snRNA since the strain expresses both the wild-type chromosomal copy of U1 and the U1-C4U allele which is expressed from a plasmid. The text describes the rationale for why the high copy expression of a factor should result in suppression of the G5A intron mutation.

transform in GAL-driven library



Select for increased copper-resistance

confirm dependence of increased copper resistance on both library and U1-C4U plasmids

Figure 2: The U1-C8U strain exhibits a loss-of-mature phenotype when splicing 5' splice site competition pre-mRNA. Primer extension analysis of RNA isolated from either a U1-wild-type strain (odd lanes) or a U1-C8U strain (even lanes). In the first 8 lanes RNA was isolated from strains transformed with a ACT1-CUP1 competition reporter, while in lanes 9 and 10, RNA was isolated from strains transformed with a non-competition ACT1-CUP1 reporter. Each lane is labeled with the specific genotype of the strain (U1 allele and intron reporter construct) from which the RNA was isolated. In all cases, the upstream 5' splice site contains the mutation and the downstream 5' splice site is wild-type. The products of the primer extension assays are indicated.

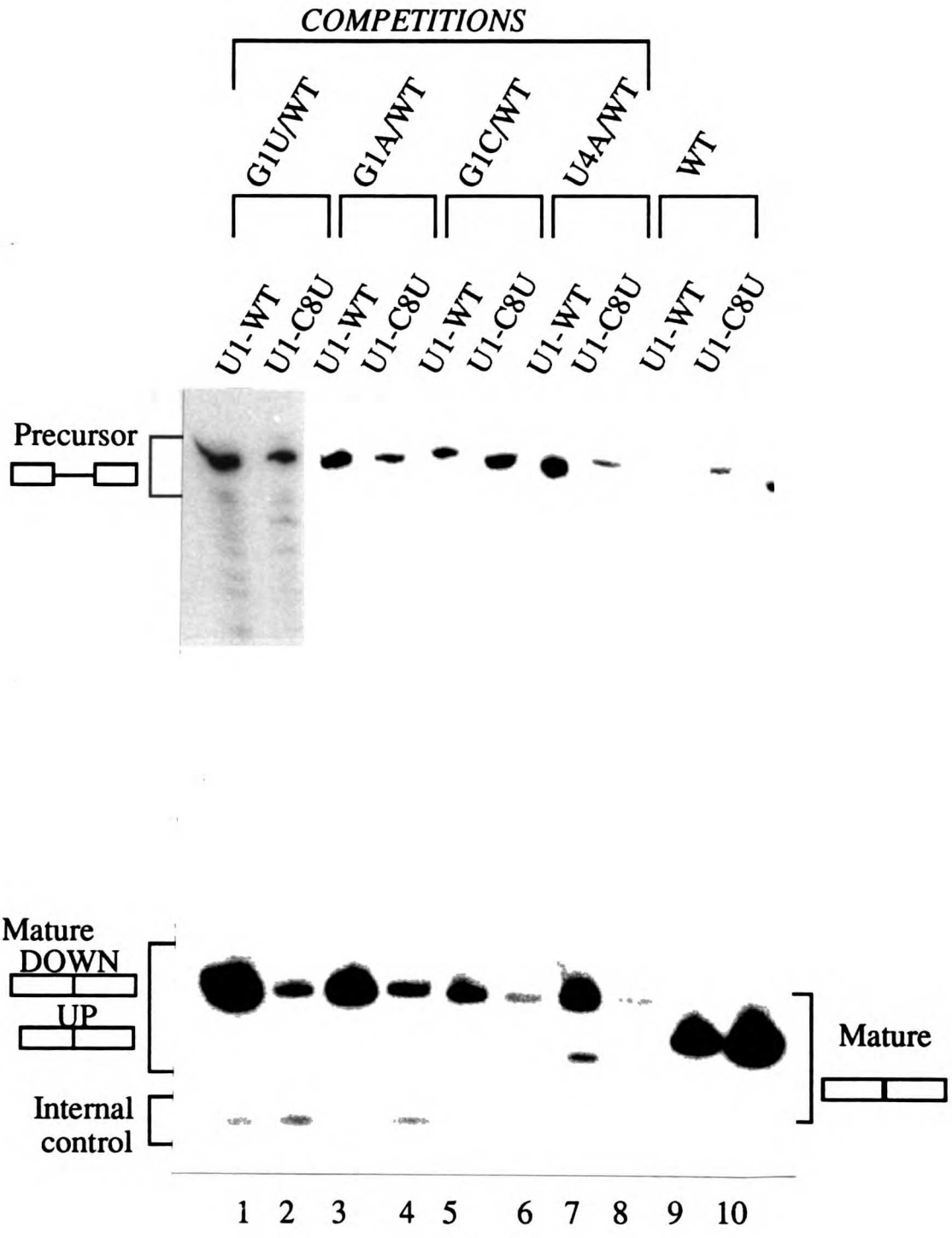

COMPETITIONS
G1U/WT
G1A/WT
G1C/WT
U4A/WT
WT
U1-WT
U1-C8U
Precursor
Mature
DOWN
UP
Internal
control
Mature
1 2 3 4 5 6 7 8 9 10

Figure 3. Suppression of intron position 4 mutations by compensating changes at U1 position 5. Primer extension analysis of RNA isolated from U1-wild-type (U1-U5) or U1-position 5 mutations (U1-U5A, U1-U5G, U1-U5C) strains carrying intron position 4/wild-type competition ACT1-CUP1 reporters. Each lane is labeled with the specific genotype of the strain (U1 allele and intron reporter construct) from which the RNA was isolated. The products of the primer extension assays are indicated.

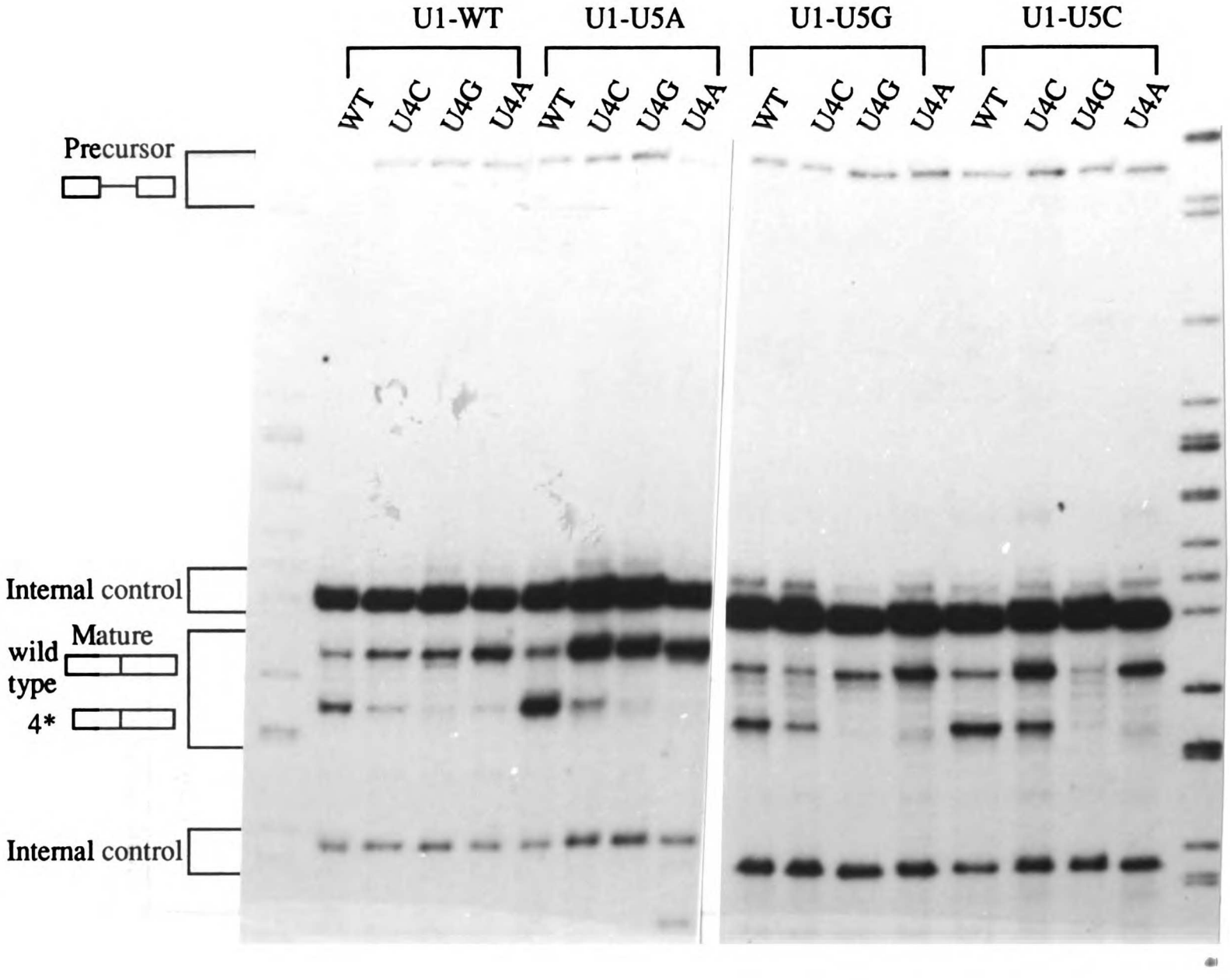

U1-WT
U1-U5A
U1-U5G
U1-U5C
WT
U4C
U4G
U4A
Precursor
Internal control
Mature
wild type
4*
Internal control

TABLE 1 Quanitation of primer extension analyses of RNA isolated from strains homozygous for U1.

intron mutation	primer extension analysis		
G5A		U1-WT	U1-C4U
	precursor	28%	9%
	lariat intermediate	21%	34%
	mature	51%	56%
G5C		U1-WT	U1-C4U
	precursor	17%	17%
	lariat intermediate	11%	11%
	mature	72%	72%
		U1-WT	U1-C8U
G1A	precursor	41%	25%
	lariat intermediate	59%	75%
	mature	0%	0%
G1C		U1-WT	U1-C8U
	precursor	85%	72%
	lariat intermediate	15%	28%
	mature	0%	0%

Table 2 Summary of growth phenotypes conferred by U1 snRNA mutations.

	Doubling time (Hrs)	Relative growth rates			
	30°C	18°C	25C	30°C	37°C
U1-WT	2.1	+++	+++	+++	+++
U1-U5C	2.8	+++	+++	+++	+++
U1-U5A	3.4	++	++	++	++
U1-U5G	3.6	++	++	++	++
U1-C4U	4.0	-	+	++	+
U1-C8U	6.4	-	-	++	+

Table 3 Summary of amount of cleavage at upstream site relative to cleavage at the downstream site.*

	mRNA upstream site/mRNA downstream site			
	U4U/WT	U4C/WT	U4G/WT	U4A/WT
U1-U5WT	3.0	.25	.08	**.06**
U1-U5C	1.7	.34	**.30**	.03
U1-U5G	1.9	**1.6**	.30	.03
U1-U5A	**4.0**	.14	.03	.01

* The bold underlined values indicate the greatest value for each of the intron mutations.

CHAPTER 4

Determinants of 5' splice site selection

Abstract

In mammals, the context of a 5' splice site as well as the proximity of a 5' splice site to the 3' splice site can influence 5' splice site selection. The following studies indicate that the context and position of the 5' splice site also play such a role in *S. cerevisiae*. In mammals, proteins are known to play an important role in distinguishing 5' splice site context and position. Though similar proteins in yeast have not been identified, the following studies indicate that some factor (or factors), in addition to U1 snRNA, plays a role in selecting 5' splice sites in yeast.

Introduction

Mammalian 5' splice site competition assays have proven to be powerful tools for analyzing the 5' splice site consensus sequence and for identifying potential factors involved in 5' splice site selection. When mammalian 5' splice site sites are in close proximity ("near competition"), splicing occurs more often at the site with greater complementarity to U1 (Eperon, et al., 1986; Nelson and Green, 1988). Even if a 5' splice site is perfectly complementary to U1, however, cleavage can occur at this site only if it is in an appropriate position and context in the pre-mRNA (Nelson and Green, 1988). This observation provided the first hint that factors other than U1 are involved in 5' splice site selection.

When competing 5' splice sites are not in close proximity ("far competition"), and the 5' splice site sequences are identical, both the sequence flanking the splice sites (exonic sequences) and the proximity of the 5' splice sites to the 3' splice site affect 5' splice

site selection (Reed and Maniatis, 1986). Thus, in mammals, factors appear to play roles in both near and far 5' splice site competitions. Interestingly, the use of competing splice sites could be altered by diluting splicing extracts, suggesting that such factors are limiting in certain circumstances (Reed and Maniatis, 1986).

Two classes of proteins, hnRNPs and SR proteins, have been implicated in influencing splice site selection. The relative concentrations of hnRNP A1 and the SR protein ASF/SF-2 can affect use of the competing 5' splice sites in the mammalian competition assay in which the splice sites are far apart (Mayeda and Krainer, 1992). hnRNP proteins are involved in packaging nascent RNAs and presenting them to the RNA processing apparatus. Some of the hnRNP proteins bind specifically to the polypyrimidine sequence of the pre-mRNA, which is part of the mammalian 3' splice site consensus sequence (Buvoli, et al., 1988; Gil, et al., 1991; Kumar, et al., 1987; Patton, et al., 1991; Swanson and Dreyfuss, 1988). In addition, the cross-linking of two of these proteins, hnRNP A1 and hnRNP C, to the pre-mRNA requires the presence of both intact U1 and U2 snRNPs (Mayrand and Pederson, 1990). Thus, both hnRNPs and snRNPs bind to splice site sequences and might interact to define splice sites.

The first member of the SR family identified, ASF/SF-2, was isolated in two independent sets of experiments; first, as a factor involved in alternative 5' splice site choice (Ge and Manley, 1990) and second, as an essential splicing factor (Krainer, et al., 1990). Curiously, all of the SR proteins tested can complement a splicing-deficient extract, but each protein exerts specific effects on the

splicing of competing 5' splice sites (Krainer, et al., 1990; Zahler, et al., 1993). These experiments demonstrate that an essential and therefore presumably constitutive splicing factor can also participate in the regulation of alternative splicing. In addition, at least in mammals, it appears that the process of splice site selection begins prior to U1 binding, since it has been shown that SR proteins can commit pre-mRNAs to the mammalian splicing pathway in the absence of U1 (Fu, 1993) and can promote U1 binding to the 5' splice site (Zuo and Manley, 1993).

Since the mammalian 5' splice site competition assays had proved so powerful in analyzing the mammalian 5' splice site, I decided to use similar assays to characterize the determinants of yeast 5' splice site selection. Chapter 2 described the development and characterization of a yeast 5' splice site competition assay in which the competing sites are in close proximity to one another (13 ntes.). These experiments indicated that mutation of any of the well conserved or invariant nucleotides of the intron portion of a 5' splice site severely decreases or eliminates the ability of this site to compete with a wild-type site. Chapter 3 demonstrated that in this competition assay, the ability of a 5' splice site to compete does not depend on its complementarity to the 5' end of U1 at least in the case of intron positions 1 and 5. This result was very surprising since, when intron position 1 and 5 mutations are not competing with a wild-type site, compensating mutations in U1 increase cleavage at the mutant 5' splice site. In addition, in the case of a second yeast 5' splice site competition assay in which the

sites are far apart (75 nt.) (Goguel, et al., 1991), cleavage at a 5' splice site is dependent on complementarity to U1.

A major difference between my work (Chapter 3) and that of Goguel et al. (1991) is the distance between the sites. Because my results suggested that as in mammals the process of 5' splice site selection involved factors in addition to U1, I wanted to further characterize the determinants of yeast 5' splice site selection. I designed three additional 5' splice site competition assays in which the competing sites are far apart. I then analyzed the behavior of these new competition assays, in both wild-type and mutant U1 strains. The results demonstrate that in yeast, as in mammals, both splice site context and proximity can play important roles in 5' splice site selection. These experiments indicate that splicing components other then U1, perhaps yeast homologs of SR or hnRNP proteins, may play a role in yeast 5' splice site selection. Importantly, the detailed characterization of the different competition assays allows several predictions as to the nature of these factors. This information can now be used to design genetic screens to identify and characterize proteins involved in yeast 5' splice site selection.

RESULTS

The context of the 5' splice site is important. Two 5' splice site competition assays (II and III) were constructed (see Figure 1) to address the question of whether the sequence flanking the 5' splice site influences splice site selection. In both of these assays

the splice sites are separated by slightly greater than 100 nucleotides. In the first assay (II) the distal (upstream) 5' splice site is flanked by synthetic sequence and the proximal (downstream) 5' splice site is flanked by authentic actin sequence while in the second assay (III) both the distal and proximal 5' splice sites are flanked by authentic actin sequence. Figure 2 aligns the synthetic sequence flanking the distal 5' splice site in competition II with the authentic sequence flanking the distal 5' splice site in competition III. The synthetic sequence is identical to the sequence flanking the distal 5' splice site in the original "near competition" assay described in Chapter 2.

In order to test if context affects 5' splice site choice, competition assays II and III were introduced into a U1-wild-type strain as well as the 5 viable strains which contain mutations in the 5' end of U1 (U1-C4U, U1-U5A, U1-U5G, U1-U5C and U1-C8U). RNA was isolated from each of these strains and assayed by primer extension analyses.

The data summarized in Table 1 are consistent with the conclusion that the context of the 5' splice site influences the amount of cleavage at that site. For example, when the distal site is flanked by synthetic sequence, as in competition II, cleavage at the distal site occurs 0.6 times as often as at the proximal site which is flanked by authentic sequence (Figure 3, lane 1; Table I). In contrast, when the distal site is flanked by authentic sequence, as in competition III, cleavage at the distal site occurs 1.4 times as often as at the proximal site which is flanked by authentic sequence (Figure 3, lane 7; Table 1). Thus, relative to the distal synthetic

site, the distal authentic site increases the ability of this site to compete against a proximal authentic actin 5' splice site by a factor of 2.4. In all of the mutant-U1 strains, a similar effect is observed.

Proximity of the 5' splice site to the 3' splice site influences splice site selection. In order to address the question of whether the proximity of the 5' splice site to the other splice site signals influences selection, an additional 5' splice site competition assay (IV) was constructed (Figure 1). In this assay, the distal site is flanked by authentic actin sequence while the proximal site is flanked by synthetic sequence. The splicing of this construct can be compared to competition assay II in which the distal site is flanked by synthetic actin sequence and the proximal site is flanked by authentic actin sequence (Figure 1). Thus, the only difference is the relative positioning of the authentic and synthetic sequences. In a wild-type U1 strain, a 5' splice site that is flanked by authentic sequence is always favored regardless of whether the 5' splice site is located in the distal site or the proximal site (Figure 4, lane 1 vs. lane 7, Table 2). The same is true in all of the mutant-U1 strains (Figure 4, Table 2). Thus, the sequence context of a 5' splice site has a dominating influence on whether a site will be used.

In the constructs used, cleavage at the proximal site results in an in-frame mRNA while cleavage at the distal site results in an out-of-frame mRNA. Thus, the ratio of cleavage of the distal site to cleavage at the proximal site is skewed in favor of the proximal site (see Chapter 2). Thus, it is complicated to determine, by comparing

these constructs, whether the proximal authentic site or the distal authentic site is preferred more.

Interestingly, in the mutant U1 strains, the amount of cleavage at either distal site relative to their respective proximal site is almost always decreased with respect to the wild-type U1 strain (Figure 4, Table 2). Thus, even if the distal site is flanked by authentic sequence, the mutant U1 strains do not recognize the distal site as well as the wild-type U1 strain. Still, a comparison of competition assays II and IV does indicate that the distal site can compete better against the proximal site when it is flanked by authentic sequence. Thus, it is apparent that even in yeast, in mutant U1 strains, both the context and proximity of a 5' splice site to the 3' splice site can influence 5' splice site choice.

Is there a limiting factor that can alter cleavage at competing 5' splice sites? The conclusion that both sequence context and splice site proximity can affect 5' splice site choice is reminiscent of previous conclusions drawn from analyses of in vitro 5' splice site competition assays in mammalian splicing extracts (Reed and Maniatis, 1986). As mentioned in the introduction, in higher eukaryotic in vitro splicing extracts, increasing the amount of exogenous added SR proteins can alter the amount of cleavage at one competing splice site relative to another in a dosage-dependent manner (Krainer, et al., 1990; Zahler, et al., 1993). If similar proteins exist in yeast, the relative amount of cleavage at two competing 5' splice sites may be affected by the concentration of such a protein in the cell. For example, one reason why synthetic

splice sites are not able to compete as well as authentic sites might be because the synthetic site cannot compete as well for a limiting protein. Thus, if the protein is limiting because the 5' splice site competition assays are expressed at high levels, it might be possible to see an increase in cleavage at the synthetic sites by decreasing the amount of pre-mRNA in the cell.

To test this prediction, competition assays II and III were subcloned from high copy (2μ) plasmids to low copy (cen) plasmids. These low copy constructs were then introduced into the wild-type and mutant-U1 strains. Preliminary experiments indicate that, in the mutant U1 strains, lowering the amount of competition RNA in the cell increases cleavage at the distal synthetic site relative to the proximal authentic site in competition assay II (Table 3), as predicted. Decreasing the copy number of the competition II has little effect on the relative amount of cleavage at the two competing authentic 5' splice sites (Table 2).

These preliminary experiments raise the exciting possibility that the in vivo concentration of a yeast splicing factor can affect 5' splice site choice. Thus, one could use the competition assay reporters to design genetic screens to identify factors which can affect 5' splice site selection.

Discussion

Analyses of mammalian 5' splice site competition assays have demonstrated that the process of 5' splice site selection is complex

and is likely to involve several other factors in addition to U1. One explanation for this complexity is that 5' splice site competition assays reflect situations of regulated splicing in which factors other than U1 act. For example, ASF/SF-2, the first identified SR protein, was independently identified as an essential splicing protein (Ge and Manley, 1990) and as a regulator of alternative 5' splice site selection (Krainer, et al., 1990). In *S. cerevisiae*, however, no examples of alternative splicing have yet been identified. Thus, I did not expect 5' splice site selection in yeast to be influenced, as it is in mammals, by both the sequence flanking the splice site, the context, and the location of the 5' splice site from the branch point and 3' splice site sequences.

Factor X is involved in recognition of the 5' splice site. In Chapter 2, I described genetic evidence which supports the hypothesis that a factor X in addition to U1 snRNA interacts with the 5' splice site consensus sequence. Factor X appears to bind preferentially to wild-type 5' splice site sequences and helps mutant U1 snRNAs recognize or bind to a 5' splice site. Below, as the result of the analyses of three additional "far competition" assays, I propose that two other factors influence the 5' splice site selection. These factors, unlike factor X, do not distinguish between sites as a result of differences in their 5' splice site sequences, but rather distinguish between sites by comparing their flanking sequence and relative position in the pre-mRNA. Curiously, each of these proposed factors, X, Y and Z, are sensitive to parameters of 5'

splice site recognition which have been shown to be important in 5' splice site selection by SR proteins.

Factor Y recognizes authentic 5' splice sites. A comparison of the splicing patterns of a set of 5' splice site competition assays (II, III and IV) in which wild-type 5' splice sites were flanked by either authentic or synthetic actin sequence, resulted in the conclusion that in yeast, like mammals, the context of a 5' splice site does influence selection of the site. What "reads" this context? Imagine that there is a factor, factor Y, that helps U1 bind to a 5' splice site and prefers a specific 5' splice site because of its context. Thus, when both competing 5' splice sites are flanked by authentic actin sequence (competition III), factor Y could bind equally to each 5' splice site and the amount of cleavage at both sites would be similar. However, when one site is flanked by authentic sequence and the other by synthetic (competitions II and IV), the synthetic cannot compete as well for factor Y as the authentic site. Thus, cleavage is favored at the site flanked by authentic sequence.

Factor Z activates distal 5' splice sites. In competition assays II, III and IV, a comparison of the wild-type U1 and mutant U1 snRNA splicing patterns demonstrated that, in general, mutant U1 snRNAs favor use of the proximal competing 5' splice sites. These observations can be explained if there is a factor, factor Z, which activates distal competing 5' splice sites but is dependent on wild-

type U1. Thus, mutant-U1 snRNAs cannot recognize or bind factor Z, and are impaired for cleavage at distal 5' splice sites.

In competition III, for example, both splice sites are flanked by authentic sequence and presumably bind factor Y. Thus, any differences in the relative amount of cleavage at the competing sites is due to something other than factor Y. In general, the mutant U1 snRNAs, when compared to the wild-type U1 strain, decrease cleavage at the distal site relative to the proximal site. Presumably, these mutants cannot interact with factor Z and are impaired for recognition of the distal site.

Variables that influence 5' splice site selection. Three variables, splice site sequence, context and location, have now been implicated in affecting 5' splice site selection. The easiest way to think about how these variables interact to select the 5' splice site is to imagine that each variable is monitored by a specific factor, thus factors X, Y and Z have been hypothesized to influence 5' splice site selection. However, there is no evidence that rules out the possibility that one factor actually monitors more than one variable. As shown in Figure 5A, factors X and Y might first bind to the pre-mRNA and then recruit U1 to the pre-mRNA. The binding of X might serve to ensure that the sequence with greatest identity to the 5' splice site is chosen while the binding of Y ensures that the 5' splice site is located in the correct region of the pre-mRNA by examining the sequence flanking the 5' splice site. The binding of these factors "recruits" U1 snRNA to the 5' splice site. In this scenario, it appears that factors X and Y might act in an analogous fashion to SR proteins

since, as discussed earlier, SR proteins have been demonstrated to recognize 5' splice sites in a context dependent manner in vitro (Krainer, et al., 1990) and in addition have been demonstrated to mediate U1 binding to the 5' splice site in vitro (Zuo and Manley, 1993).

Factor Z appears to favor cleavage at distal 5' splice sites. Thus, in *S. cerevisiae*, where no examples of alternative splicing have been observed, it is not clear what role such a protein might play in normal 5' splice site selection. One interesting possibility is that factor Z coordinates the interaction between U1, the 5' splice site and branch point sequence, an interaction that is essential for committing a pre-mRNA to the splicing pathway (Legrain, et al., 1988; Seraphin and Rosbash, 1989b), in a fashion analogous to the mammalian U2AF protein. Thus, in the competition assay, factor Z could be responsible for ensuring that the branch point sequence is intact before committing U1 to a particular 5' splice site regardless of its proximal or distal location (Figure 5B). However, in the absence of factor, perhaps cleavage is favored at the proximal site because the 5' splice site which is in closer proximity to the branch point recognition sequence is more likely to interact with this sequence by random collision. Thus, in the case of the U1 mutants, which are hypothesized to be unable to bind Z, cleavage is diminished at the distal site because the proximal 5' splice site preferentially interacts with the nearby branch point sequence (Figure 5C). In other words, mutant U1 snRNAs are diminished for cleavage at the upstream site because they lack factor Z and are unable to interact with the branch point recognition sequence

efficiently and thus commit the distal 5' splice site to the splicing pathway (Figure 5C).

Can U1 compensating mutations suppress 5' splice site mutations in the competition assay? In Chapter 2, I described that the U1-C4U mutation was not able to suppress a G5A 5' splice site mutant in the "near competition," even though suppression of this specific mutation was demonstrated in a different "far competition" assay (Goguel, et al., 1991). Thus, the initial incentive for increasing the distance between the competing 5' splice sites was to see if it the spacing between the competing 5' splice sites was inhibiting suppression. I have now introduced the G5A mutation into the proximal authentic site of "far" competition II. The strong bias of the U1-C4U mutant strain for the proximal site complicates the interpretation of the data. Consequently, it is not clear if suppression can be detected. Additional experiments will need to be conducted to resolve this issue.

In summary. The process of 5' splice site selection in *S. cerevisiae* , like that of mammals, appears to involve other factors in addition to U1 snRNA. The 5' splice site sequence, context and positioning in the pre-mRNA all influence the ability of a splice site to be recognized. In higher eucaryotes pre-mRNA splicing often serves as a mechanism for regulating gene expression. Thus, one rationale for the complexity of 5' splice site selection is that this early step in the splicing pathway can be exploited for regulation. However, in yeast, no examples of alternative splicing and only two

examples of regulated splicing have been observed (Dabeva, et al., 1986; Engebrecht, et al., 1991). Thus, on one hand, it is surprising that the process of 5' splice site selection would prove to be so complicated in yeast. Although, the complexity of the process of 5' splice site selection might serve to ensure the fidelity of pre-mRNA recognition. On the other hand, a common mechanism of regulating splicing is by exploiting the roles of normal splicing factors. So perhaps, the yeast competition assays indicate how it is possible to modify the normal splicing apparatus to regulate pre-mRNA splicing. Thus, it is likely that other examples of regulated splicing will be found in yeast.

Figure 1. Schematic representation of competition assays II, III and IV. The open boxes represent exons and the thin line introns.

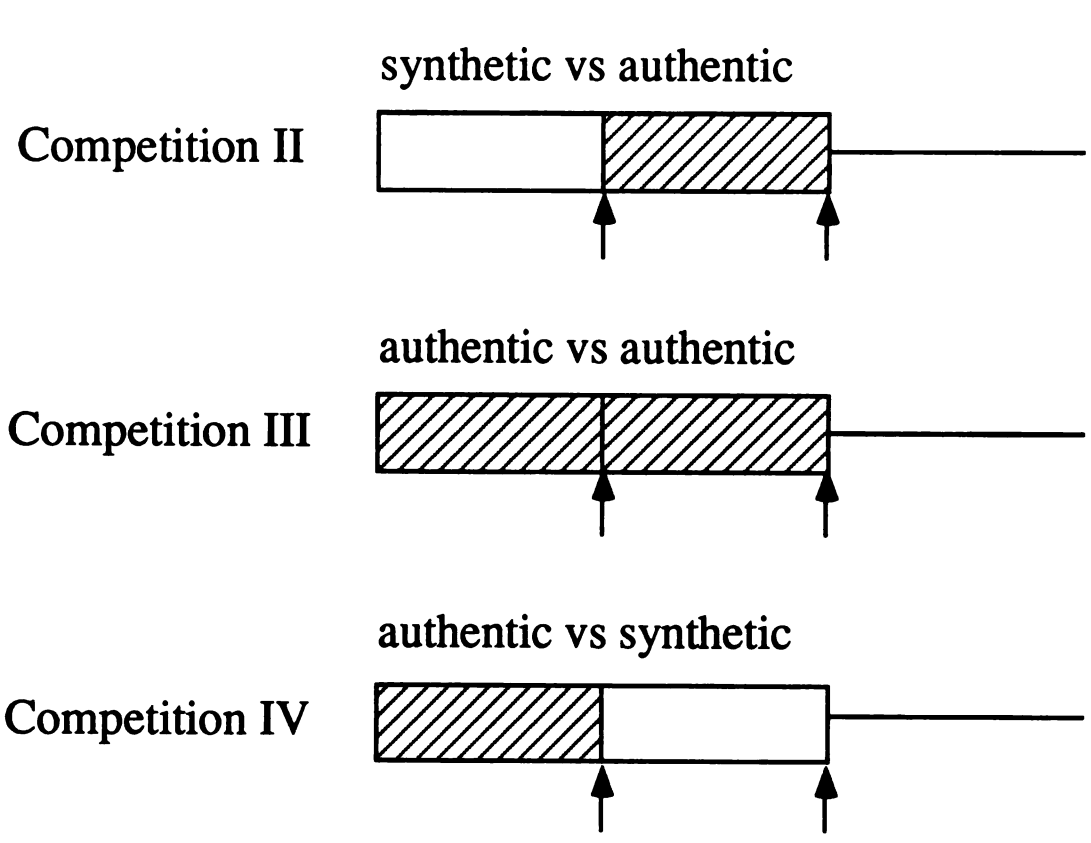

DISTAL
PROXIMAL
synthetic vs authentic
Competition II
authentic vs authentic
Competition III
authentic vs synthetic
Competition IV
= authentic context
= synthetic context
= cleavage site

Figure 2. Comparison of the sequences flanking the distal 5' splice sites in competitions II and III. The sequence flanking competition II is synthetic actin sequence while the sequence flanking competition III is authentic actin sequence. The bold sequence is the actual 5' splice site sequence.

competition II

AAAAUUUACUGAAUGAGAUCU ***AUG GUAUGU*** UUAAAUGAGGUUGCGGGUU

AAAAUUUACUGAAUUAACAAUGGAUU ***CUG GUAUGU*** UCUAGCGCUUGCG

competition III

↑ = cleavage site

Figure 3. A comparison of the splicing patterns of competitions II and III. Primer extension analysis of RNA isolated from strains homozygous for wild-type or mutant U1 snRNA transformed with either competition II or competition III ACT1-CUP1 reporters. Each lane is labeled with the specific U1 strain and the competition reporter from which the RNA was isolated and the products of the primer extension assays are indicated on the side of the figure.

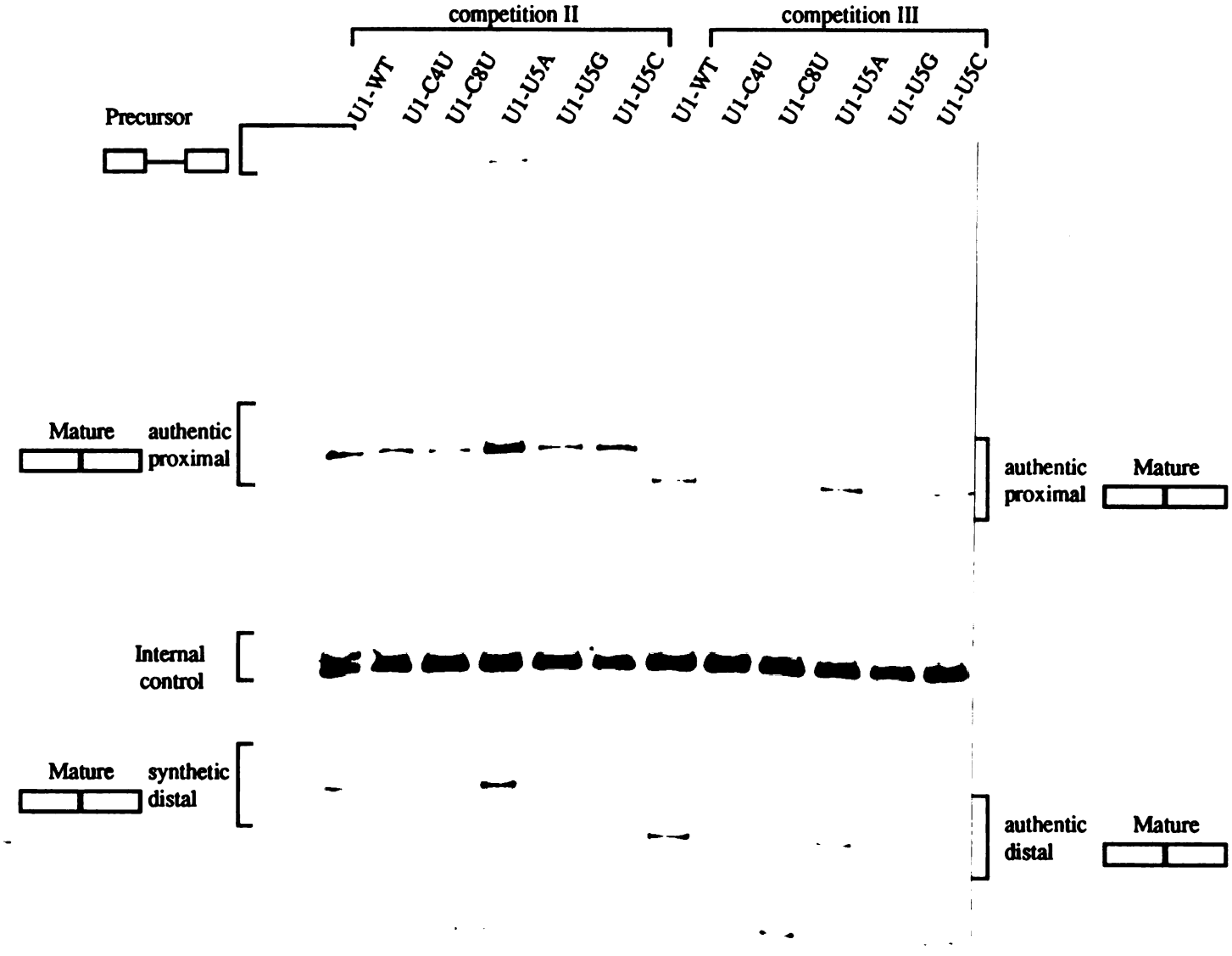

competition II
competition III
U1-WT
U1-C4U
U1-C8U
U1-U5A
U1-U5G
U1-U5C
U1-WT
U1-C4U
U1-C8U
U1-U5A
U1-U5G
U1-U5C
Precursor
Mature
authentic
proximal
authentic
proximal
Mature
Internal
control
Mature
synthetic
distal
authentic
distal
Mature

Figure 4. A comparison of the splicing patterns of competitions II and IV. Primer extension analysis of RNA isolated from strains homozygous for wild-type or mutant U1 snRNA transformed with either competition II or competition IV ACT1-CUP1 reporters. Each lane is labeled with the specific U1 strain and the competition reporter from which the RNA was isolated and the products of the primer extension assays are indicated on the side of the figure.

In these competitions, in contrast to the ones described in Figure 2, cleavage at the proximal site results in the production of an in-frame ACT1-CUP1 reporter gene and cleavage at the distal site results in the production of an out-of-frame ACT1-CUP1 reporter gene. Note, a comparison of the first lanes of figures 1 and 2 exemplifies how the relative reading frames of the mRNAs can affect the interpretation of the data. In Figure 2, where cleavage at the distal site results in the production of an in-frame mRNA, it appears that the proximal site is used 1.6 times as often as the distal site. In contrast in Figure 3, where cleavage at the proximal site results in the production of an in-frame mRNA, it appears that the proximal site is used 8.3 times as often as the distal site. These results demonstrate how important it is to consider the relative stability of the mRNA produced when comparing two competition constructs.

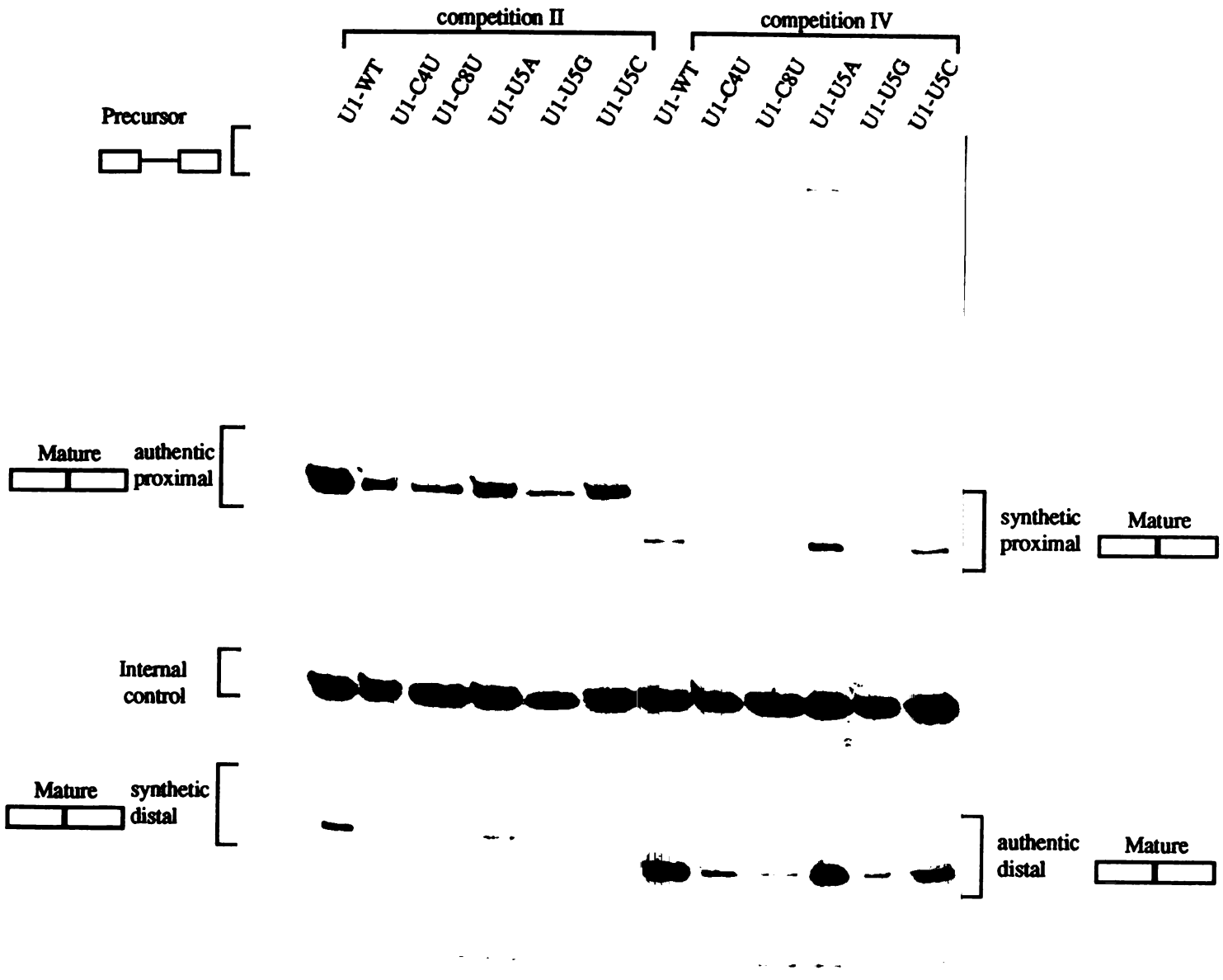
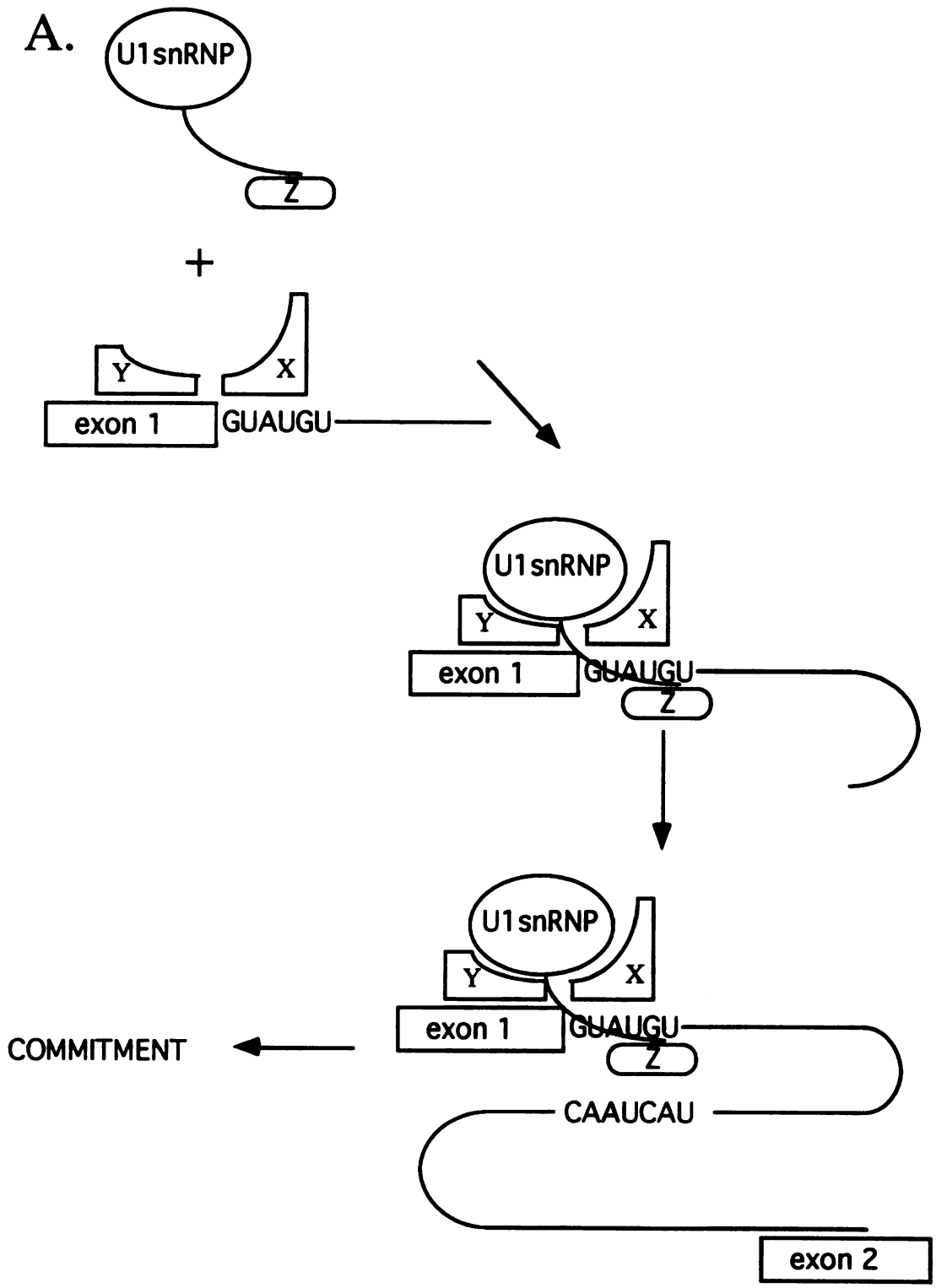
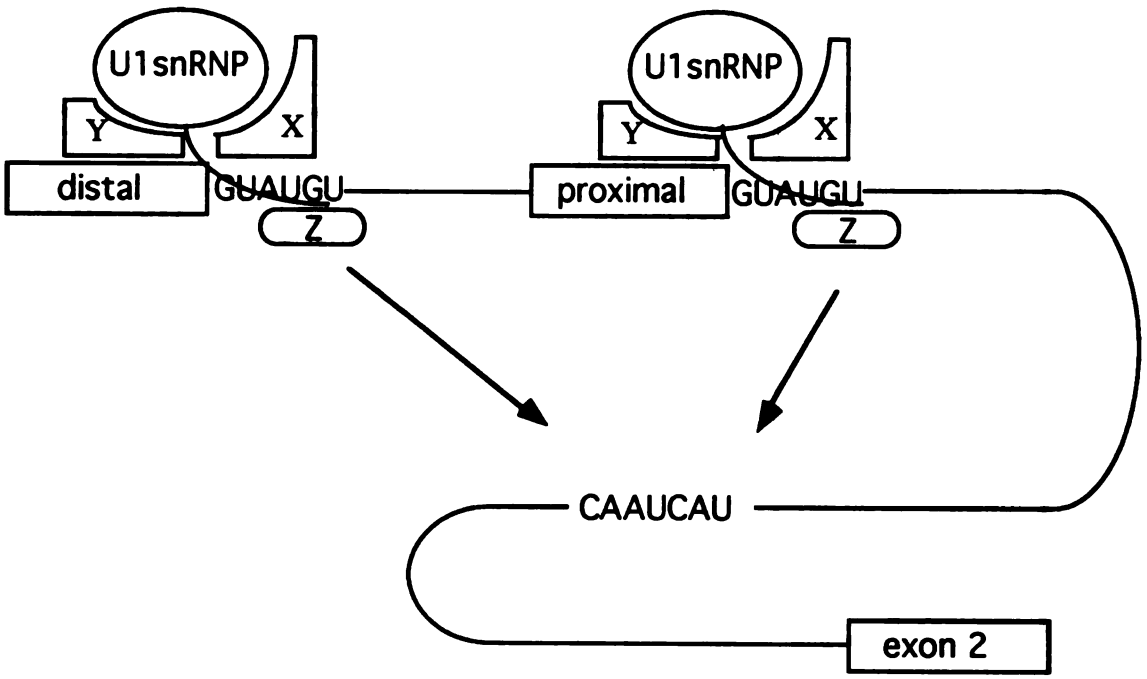

competition II
competition IV
U1-WT
U1-C4U
U1-C8U
U1-U5A
U1-U5G
U1-U5C
U1-WT
U1-C4U
U1-C8U
U1-U5A
U1-U5G
U1-U5C
Precursor
Mature
authentic
proximal
synthetic
proximal
Mature
Internal
control
Mature
synthetic
distal
authentic
distal
Mature

Figure 5. Model for the roles of factors X, Y and Z. **(A)** Depicted in the situation that occurs in a normal non-competition situation. **(B)** Shown is what is hypothesized to happen in a "far" competition assay with a wild-type U1. **(C)** Shown is what is hypothesized to happen in a "far" competition assay with a U1 mutated at its 5' end. (see text for details of the models)


A.
U1snRNP
Z
+
Y
X
exon 1
GUAUGU
U1snRNP
Y
X
exon 1
GUAUGU
Z
U1snRNP
Y
X
exon 1
GUAUGU
Z
COMMITMENT
CAAUCAU
exon 2

B.


U1snRNP
Y
X
distal
GUAUGU
Z
U1snRNP
Y
X
proximal
GUAUGU
Z
CAAUCAU
exon 2

C.

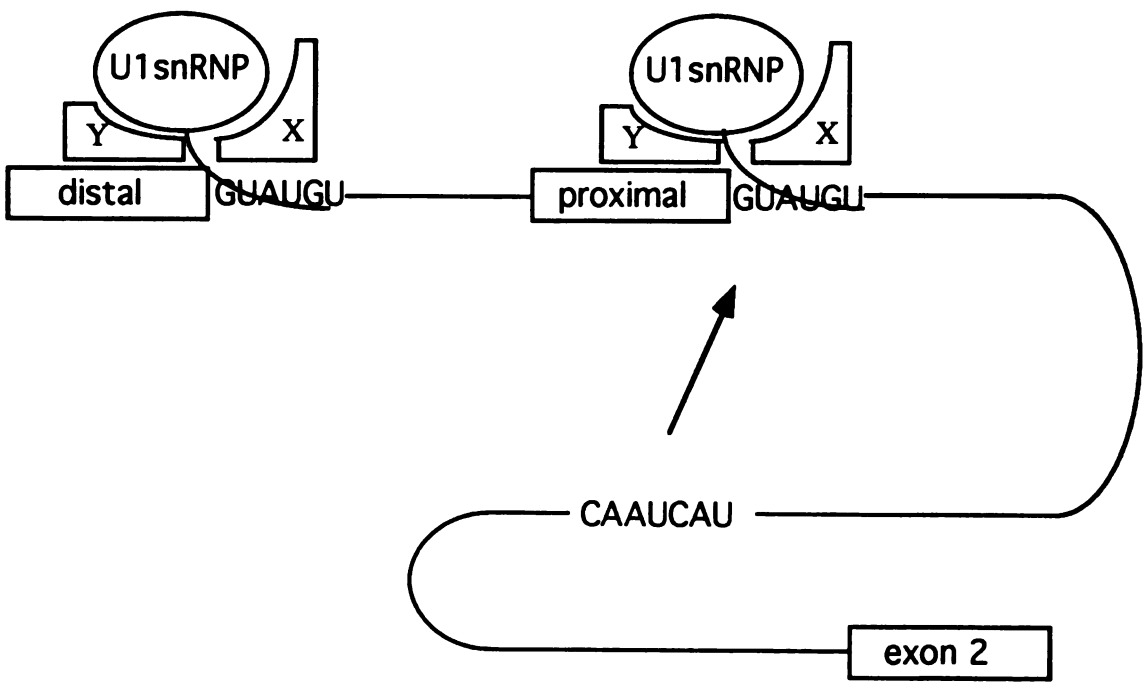

U1snRNP
Y
X
distal
GUAUGU
U1snRNP
Y
X
proximal
GUAUGU
CAAUCAU
exon 2

Table 1 A comparison of the ability of synthetic and authentic proximal 5' splice sites to compete with a proximal authentic 5' splice site.

distal/ proximal	comp. II **synthetic*†**/ authentic	comp. III **authentic**/ authentic	comp. III/ comp. II
U1-wild type	.63	1.4	2.4
U1-U5A	.32	.77	2.4
U1-U5C	.20	.91	4.6
U1-U5G	.26	1.7	6.0
U1-C4U	.18	.59	3.2

* The bold font indicates that cleavage at this site generates a stable in-frame mRNA; otherwise cleavage at the site generates an unstable out-of-frame mRNA.

† The ratio = distally spliced/proximally spliced.

Table 2 A comparison of the ability of a proximal synthetic and a distal synthetic site to compete with an authentic site.

distal/ proximal	comp. II synthetic†/ **authentic***	comp. IV authentic/ **synthetic**
U1-wild type	.12	7.1
U1-U5A	.14	3.0
U1-U5C	.06	4.2
U1-U5G	.06	2.4
U1-C4U	.07	2.9

* The bold font indicates that cleavage at this site generates a stable in-frame mRNA; otherwise cleavage at the site generates an unstable out-of-frame mRNA.

† The ratio = distally spliced/proximally spliced.

Table 3 A comparison of high and low copy expression of competitions II and III.

	competition II **synthetic***† /authentic			competition III **authentic**/synthetic		
copy #	high	low	low/hi	high	low	low/hi
U1-WT	.63	.63	1.0	1.4	1.1	.79
U1-U5A	.32	.83	2.6	.77	1.1	1.4
U1-U5C	.20	.50	2.5	.91	.77	.85
U1-U5G	.26	.56	2.1	1.7	2.0	1.2
U1-C4U	.18	.45	2.5	.59	1.0	1.7

* The bold font indicates that cleavage at this site generates a stable in-frame mRNA; otherwise cleavage at the site generates an unstable out-of-frame mRNA.

† The ratio = distally spliced/proximally spliced.

CHAPTER 5

Mutations in U6 snRNA that Alter Splice Site Specificity: Implications for the Active Site

ABSTRACT

Mutation of the invariant G at position 5 of 5' splice sites in introns of *S.cerevisiae* (/GUAUGU) activates cleavage at nearby aberrant sites. Using a sensitive in vivo reporter assay, we have tested a proposed base-pairing interaction between the 5' splice site and the invariant ACAGAG sequence of U6 that is mutually exclusive with the U1/5' splice site interaction. Mutations in U6 and/or the intron predicted to stabilize this interaction suppressed aberrant cleavage and increased normal cleavage. In contrast, aberrant cleavage was enhanced by stabilizing an alternative base-pairing interaction between the intron and a non-conserved U6 sequence upstream of ACAGAG. We propose that base-pairs between the intron (/GUAUGU) and U6 (ACAGAG) play an important role in aligning the 5' splice site with the branch point during the first catalytic step. In addition, nearby mutations in U6 can also influence 5' cleavage. Finally, we found that mutation of the last nucleotide of the ACAGAG sequence can suppress the second step block imposed by mutations of the 3' splice site dinucleotide (AG/). These data can be accommodated by a model for the spliceosomal active site which juxtaposes the branch point and 5' and 3' splice sites via a set of RNA:RNA interactions between the intron, U2 and U6.

Nuclear pre-mRNA splicing is the process by which introns are removed from mRNA precursors to form functional mature mRNAs. This process is mediated by the spliceosome, a complex ribonucleoprotein machine composed of five small nuclear RNAs (U1,U2,U4,U5 and U6) and numerous proteins. Introns are removed via two consecutive transesterification reactions. In step 1, the 2'-OH of the intron branch point adenosine attacks the phosphodiester bond at the 5' splice site, resulting in the formation of free exon 1 and the lariat intermediate. In step 2, the 3'-OH of exon 1 attacks the 3' splice site, resulting in ligation of exon 1 to exon 2 and the release of the excised intron lariat (*1,2*). Introns are only defined by three short consensus sequences: the 5' splice site, the branch point region, and the 3' splice site. It is not understood how the spliceosome correctly identifies boundaries between exons and introns and chooses the precise location of the sites of covalent modification.

There is now considerable information on the role of RNA-RNA interactions in this identification process. The 5' splice site and branch point sequences of the intron are recognized by base-pairing interactions with U1 (Fig. 1A)(*3,4*) and U2 (*5*), respectively. U1 binds at an early step in spliceosome assembly and commits the pre-mRNA to the splicing pathway (*6*). Genetic and biochemical data place U5 in close proximity to the 5' and 3' exon sequences (*7-10*). In special cases, in which the authentic 5' splice site has been mutated, Watson-Crick base-pairing interactions between U5 and the 5' region of the precursor mRNA play a role in activating aberrant cleavage sites (*8*).

Mutational studies have shown that the highly conserved 5' splice site of *S. cerevisiae* plays several roles in the splicing pathway (*11*). Mutation of the invariant G at intron position 5 (/GUAUGU)(*12*), like other 5' splice site mutations, decreases the efficiency of the first catalytic step. However, changes at this position are the only intron mutations that activate aberrant cleavage sites. The location and frequency of aberrant cleavage depends on the particular intron mutated as well as the specific intron mutation (*13*). Restoration of complementarity between the 5' splice site and U1, by the introduction of compensatory changes in U1, does not prevent aberrant cleavage; rather, it increases cleavage at both the normal and aberrant sites (Fig. 1B)(*4*). Thus, while the base-pairing interaction between the invariant C at position 4 of U1 and the invariant G at position 5 of the 5' splice site is important for the efficiency of the first catalytic step of splicing, intron position 5 must also be recognized by some other factor in order to specify the location of the cleavage reaction (*4*).

An excellent candidate for such a second factor is U6. Crosslinking experiments in mammalian (*10*) and yeast (*14*) extracts indicate that intron position 5 is close to a conserved domain of U6. Two distinct U6-5' splice site base-pairing interactions have been proposed based on these experiments (*15*). In both models, ACA sequences in U6 base pair with the UGU conserved at positions 4 to 6 of yeast introns. In model I (Fig. 1C), which is based on a psoralen crosslink between U6 and pre-mRNA (*10*), the ACA residues (nt. 47 to 49) are in the invariant ACAGAG sequence. In model II (Fig. 1D), which is based on ultraviolet (UV) crosslinks between U6 and lariat

intermediates (*14*), the ACA residues (nt. 42 to 44) are located just upstream of the ACAGAG sequence and are relatively nonconserved (*16,17*). The regions of U6 modeled to base pair with the 5' splice site are immediately upstream of a helix (helix I) formed by base-pairing interactions between U6 and U2 (*18*). This helix can juxtapose the essential ACAGAG sequence of U6 (*19,20*) with the branch point interaction domain of U2 (*18*). Thus, the proposed U6-5' splice site interactions might serve to juxtapose the branch point, which is the nucleophile that attacks the 5' splice site (*21*), with the 5' cleavage site.

We designed a genetic approach to test the role of U6 in the specificity of 5' splice site cleavage. Mutations predicted to stabilize the interactions in model I suppressed aberrant cleavage and increased normal cleavage. Conversely, enhancement of the interaction in model II led to an increase in cleavage at the aberrant site activated by a mutation at intron position 5. We thus propose that base pairs between the intron (positions 4 to 6) and U6 (nt. 47 to 49) serve to correctly align the 5' splice site with the intron branch point during the first catalytic step. We also found that mutation of the last nucleotide of the ACAGAG sequence can suppress the second step block imposed by mutations of the 3' splice site dinucleotide (AG/). These data, together with other recent results (*18,22,23*), suggest a model for the catalytic center of the spliceosome.

Genetic test of the models of interactions between U6 and the intron. In order to identify potential base-pairing interactions

between U6 and the 5' splice sites, we generated strains homozygous for U6 point mutations (*24*) and transformed them with ACT1-CUP1 reporters (Fig. 2A)(*25*) that contained either a wild-type intron (Fig. 2B) or intron position 5 mutation (G5A or G5C)(*26*). We monitored the amount of cleavage at either the normal (NL) or aberrant (AB) cleavage sites by using a pair of reporters: one produces an in-frame mRNA when cleavage is at the normal site (Fig. 2C) and the other produces an in-frame mRNA when cleavage is at the aberrant site (Fig. 2D)(*27*). The ACT1-CUP1 fusion protein allows cells to grow in the presence of copper in a dosage-dependent manner. Thus, in a cup1Δ background, it is possible to monitor mRNA by monitoring copper resistance of strains harboring the reporter constructs (*28*). For example, we can detect that the G5A mutation activates more aberrant cleavage than the G5C mutation, because wild-type strains transformed with the G5A reporter for aberrant cleavage grow at .25 mM copper, but wild-type strains transformed with the G5C reporter for aberrant cleavage grow at 0.1 mM copper.

If activation of the aberrant cleavage site results from disruption of a normal base-pairing interaction between U6 and position 5 of the 5' splice site, restoration of the proposed base-pairing interaction should inhibit cleavage at the aberrant site. The most direct test of this prediction would be, in the case of model I (Fig. 1C), to determine if mutations at U6 position 48, which should restore complementarity to intron position 5, could prevent cleavage at the aberrant site. However, strains are inviable when the sole copy of U6 is mutated at position 48 (*19*). Moreover, when expressed on low-copy plasmids in the presence of the chromosomal copy of

U6, mutations at U6 position 48 did not influence the splicing phenotypes conferred by intron position 5 mutations, as determined by primer extension and copper resistance assays (*29*). This is not surprising since these U6 alleles are poorly expressed when assayed in the presence of the chromosomal copy of U6 (*19*).

An additional prediction of model I is that cleavage at the aberrant site might be suppressed by strengthening the proposed base-pairing interaction between U6 and the 5' splice site by extending the complementarity between U6 positions 47 to 49 (ACA) and intron positions 4 to 6 (UGU) to include U6 positions 45 and 46 and intron positions 7 and 8 (Fig. 3A). Thus, U6 positions 45 and 46 were mutated to A45G,U46A to create complementarity to actin intron positions U7 and C8 (Fig. 3A). If these nucleotides are in close proximity to the 5' splice site, they should be able to base pair with the intron, thereby stabilizing the normal U6-intron base-pairing interaction (model I), and inhibiting cleavage at the aberrant site (*30*).

As predicted, when either the G5C or G5A reporter was transformed into the U6-A45G,U46A strain, cleavage at the normal site was increased but cleavage at the aberrant site was decreased, as assayed by copper resistance (Fig. 3B). In fact, the copper resistance of these mutant U6 strains, transformed with reporters for normal cleavage, was as high as that of a wild-type strain transformed with a wild-type 5' splice site reporter (Fig. 3B). Primer extension analyses (Fig. 3C) confirmed that the changes in copper resistance resulted from alterations in the amount of mature mRNAs generated by cleavage at each of the two sites and indicated

that the amount of cleavage at the normal site relative to the aberrant site was at least ten times greater in the U6-A45G,U46A strain than in the U6-wild-type strain (*31*).

A second test of the U6-intron interaction was conducted by mutating the G5A intron positions 7 and 8 (U7C,C8G) (Fig. 3A). Unexpectedly, the introduction of these mutations, into a nonconserved region of the intron (*32*), inhibited splicing in either the wild-type or the U6-A45G,U46A strains. Nonetheless, the U6-A45C,U46G mutation, which was predicted to restore complementarity between U6 and the altered intron (Fig. 3A)(*30*), restored splicing efficiency and increased cleavage at the normal site and decreased cleavage at the aberrant site as assayed by copper resistance (Fig. 3B). Primer extension analyses indicated that the amount of cleavage at the normal site relative to the aberrant site was at least twenty times greater in the U6-A45C,U46G strain than in the U6-wild-type strain (*31*)(Fig. 3C). Although we can not demonstrate a direct base-pairing interaction between intron position 5 and U6, the specificity of inhibition of aberrant cleavage by extension of the region of complementarity between U6 and the 5' splice site supports the hypothesis that base-pairing is required for accurate cleavage.

The above results suggest that intron position 5 mutations activate aberrant cleavage by weakening the normal interaction with U6 at the 5' splice site. Several observations suggested that potential fortuitous base-pairing interactions between nearby sequences in U6 and the mutated intron might re-phase the interaction and thus account for the location of the aberrant site.

As pointed out earlier, the location and efficiency of cleavage at aberrant sites depends on the intron into which the position 5 mutation is introduced (*13*). The base-pairing interaction proposed in model II was based on U6 crosslinking to the yeast actin intron (*14,15*), which allows a more extended region of complementarity between U6 and the intron (Fig. 4A). Model II aligns the UGU of the 5' splice site with an ACA sequence (nt. 42 to 44) five nucleotides upstream of the ACA (nt. 47 to 49) aligned in model I (Fig. 1C), which is the distance between the normal and aberrant cleavage sites in actin. In fact, the position of the aberrant cleavage in the actin intron can be rationalized by the base-pairing interaction proposed in model II since, as in model I, the conserved C at position 48 in U6 is complementary to the G residue five nucleotides downstream of the aberrant site (Fig. 4B).

Since aberrant cleavage can be inhibited by strengthening the base-pairing interaction proposed in model I, strengthening the base-pairing interaction in model II should increase aberrant cleavage. Therefore, as a test, we transformed U6 strains homozygous for each of the position 43 mutations with the intron G5A and G5C reporters (Fig. 2)(*33*) Only the U6-C43G strain exhibited increased copper resistance relative to the wild-type strain when transformed with the G5C reporter for cleavage at the aberrant site (Table 1). This mutant strain also exhibited decreased copper resistance relative to the wild-type strain when transformed with the G5C reporter for cleavage at the normal site (Table 1). Primer extension analyses confirmed that the U6-C43G mutation resulted in increased cleavage at the aberrant site activated by the

intron G5C mutation (Fig. 5). In contrast, when splicing the G5A reporter, the U6-C43G mutation decreased cleavage at both the normal and aberrant sites (Table 1; Fig. 5). The observed allele specificity for increasing aberrant cleavage with the G5C intron is consistent with model II since U6-C43G, unlike U6-C43A and U6-C43U, should increase the stability of the alternative base-pairing interaction (Fig. 4C)(*34*).

To complete our analysis of this region of U6, all possible point mutations were introduced at positions 38 to 46 (*33*), except for A38U and A45G, because the crosslinking experiments (*10,14*) indicated that this region of U6 is close to the precursor-mRNA. Although, many mutations in this region had slight effects on splicing when combined with mutations at intron position 5 (*35*), no change other then C43G increased the absolute amount of cleavage at the aberrant site. All three strains homozygous for point mutations at U6 position 39 exhibited increased copper resistance relative to the wild-type strain when transformed with the intron position G5A reporter for normal cleavage (Table 1). Only the G39C mutation increased cleavage at the normal site with the G5C reporter and decreased cleavage at the aberrant site with both the G5A and G5C reporters (Table 1)(*36*). We do not understand how these mutations at position 39 influence cleavage site choice, although the lack of allele specificity suggests that the mechanism differs from the U6-C43G case. However, in Ascaris this location in U6 is activated for aberrant attack by the branch point when mutations are introduced into U6 upstream of the ACAGAG sequence (*37*).

The base-pairing interaction proposed in model I (Fig. 1C) closely juxtaposes the invariant ACAGAG sequence of U6 with the 5' splice site. We analyzed the phenotypes conferred by the four U6-ACAGAG mutant strains that are viable at 30°C (A47G, A47U, G50C, and G52U)(*19*) on the position 5 mutant intron reporters. None of these strains conferred a phenotype when splicing the wild-type actin intron; however, all exacerbated the splicing defect of G5A and G5C introns. The position 47 mutations decreased cleavage at both the normal and aberrant cleavage sites, A47G>A47U (Fig. 6A), a result consistent with the severity of their blocks to the first catalytic step in vitro (*20*). The U6-G50C mutation is the only mutation which almost abolishes cleavage at the normal site with both the G5C and G5A reporters, but only decreases cleavage at the aberrant site (G5A>G5C)(Fig. 6A). In addition, the G52U mutation decreased cleavage of pre-mRNA with a wild-type 5' splice site only slightly, if at all, when assayed by in vivo (*29*) or in vitro assays (*20*), but when combined with the intron position 5 mutant reporters the U6-G52U mutation almost abolished cleavage at both the normal and aberrant sites (Fig. 6A). Thus, there was a highly synergistic effect observed when the intron position 5 and the U6-G52U mutations were combined. Taken together with the severe phenotype conferred by the G50C mutation, these results are consistent with model I, in which residues 50 to 52 of U6 are most closely juxtaposed to the cleavage site, and thus likely to function in the first catalytic step.

Suppression of 3' splice site mutations. Intron position 1, like intron position 5, plays at least two roles in the splicing pathway. Position 1 mutations decrease the efficiency of the first catalytic step and result in a complete block to the second catalytic step (*38,39*). Mutations of the 3' splice site also severely inhibit the second step (*8,38*). Parker and Siliciano (*23*) have shown that specific changes at the first (G1) and last (G303) intron positions of the actin intron can suppress the block to the second catalytic step conferred by each mutation, suggesting a direct interaction between G1 and G303. Model I juxtaposes U6 position 52 close to intron position 1 (Fig. 1C), which is presumably in proximity to G303. Moreover, certain mutations in U6-A51 and U6-G52 inhibit the second step of splicing in vitro (*20*). To test whether this region of U6 is involved in the recognition of the 3' splice site AG/, we considered the possibility that if viable U6 mutations in the invariant ACAGAG sequence (*16,17,19*) could suppress the second step phenotypes conferred by 5' or 3' splice site mutations (or both). For reporters, we used actin-LacZ constructs carrying mutations at intron position G1 or in the 3' splice site (A302,G303) or both (*23,40*). The U6-G52U mutation was able to weakly suppress the block to the second catalytic step conferred by all three point mutations at G303 (Table 2); in addition, we observed suppression of A302U, but not A302C. In contrast to this suppression with G52U, the U6-G50C and U6-A47G mutations exacerbated the phenotype conferred by the 3' splice site mutations as seen by a decrease in β-galactosidase (G50C>>A47G) (Table 2). Primer extension analyses of RNA isolated from these U6 mutants, transformed with the G303C

ACT1-CUP1 reporter (*23,41*), confirm that suppression conferred by the U6-G52U mutation was due to increased cleavage at the mutated 3' splice site (Fig. 6B)(*42*) .

Both the U6-G52U and the intron G1A mutations suppress the block to the second step conferred by the intron G303C mutation. Thus, it is possible that all three nucleotides interact during the second catalytic step of splicing. However, the U6-G52U mutation is not able to suppress the 3' splice site mutations when combined with an intron position 1 mutation (Table 1). Primer extension analyses of RNA isolated from strains transformed with the G1A-G303C ACT1-CUP1 reporter (*41*) revealed that the loss of suppression was due to a decrease in the efficiency of the first catalytic step, as demonstrated by a decrease in the accumulation of lariat intermediate (*29*). Thus, it is not possible to genetically evaluate if U6 position 52 and actin intron positions 1 and 303 interact simultaneously during the second catalytic step. Nonetheless, mutations at position 52 decrease the splicing efficiency of introns mutated at position 1 (*29,43*) or position 5 (Fig. 6A), yet increase splicing of introns mutated at the 3' splice site (Fig. 6B). These results imply that U6 position 52 plays different roles during each catalytic step of splicing.

Function of 5' splice site-U6 base-pairing interaction. As shown above, the specificity of 5' splice site cleavage can be enhanced or reduced by mutations predicted to alter two distinct base-pairing interactions between U6 and the 5' splice site in the actin intron in *S. cerevisiae*. When invariant position G5 of the

intron is altered, mutations predicted to stabilize the interaction shown in model I suppress aberrant cleavage and increase cleavage at the normal site. Conversely, enhancement of the interaction drawn in model II leads to an increase in cleavage at the aberrant site. We propose that the pairing between the conserved UGU at intron positions 4 to 6 (in *S. cerevisiae*) and the invariant ACA at positions 47 to 49 of U6 is a determinant of the proper position of cleavage at the 5' splice site. This conclusion is consistent with the observation that aberrant cleavage activated by position 5 mutants in the yeast rp51A intron was partially suppressed by compensatory changes in U6 position 48 (*44*) .

In yeast, only mutations in the invariant G at position 5 activate aberrant cleavages (*11*). Thus the G·C base pair between intron position 5 and C48 of U6 is likely to play a particularly important role in maintaining the fidelity of cleavage site choice. Although 5' splice site sequences in mammals are considerably more degenerate, position 5 is a G in 85 percent of mammalian introns (*1*). In addition, the base-pairing interaction in model I (Fig. 1C) provides an explanation for the conservation of a U at position 4 of yeast introns (*11*), even though this would prevent base-pairing with the invariant U in U1 (Fig. 1A); thus yeast can form three consecutive base pairs between U6 and the 5' splice site. Conversely, in mammals this potential interaction would be shortened to two base pairs because the consensus mammalian intron contains an A at position 4 and thereby would be increased in complementarity to U1.

Our experiments also provide an explanation for the location of the aberrant cleavage site activated by actin position 5 mutations

(*45*). Presumably, the 5' splice site region of actin can fortuitously form an extended base-pairing interaction with nonconserved sequences in U6 upstream of the invariant ACA. This alternative base-pairing mode would be favored by disruption of the normal base-pairing interaction in position 5 mutants and would be further stabilized by the creation of a G·C base pair between intron position G5C and U6-C43G. This alternative pairing would juxtapose U6-48C with G1 instead of G5 (Fig. 4). Thus in both normal and aberrant cases, cleavage would occur five nucleotides upstream of the base pair formed with C48 of U6. Although the alternative base-pairing interaction between G5C and the U6-C43G mutant is similar to that proposed for an interaction between U6 and wild-type introns (model II), the crosslinks on which this proposal was based were identified only in the lariat intermediate (*14*); thus model II may reflect a valid U6-5' splice site interaction that occurs at a different step in the splicing pathway

Other factors are also likely to contribute to the fidelity of 5' cleavage. For example, an alternate base-pairing scheme similar to that in model II cannot account for the location of the aberrant cleavage site in position 5 mutants of the rp51A intron. Mutations in the invariant loop of U5 activate aberrant cleavage when the intron is mutated at position 1 by base pairing with exon sequences adjacent to the newly activated aberrant 5' splice site (*8*). Also, the sequences just upstream of the aberrant cleavage sites activated by intron position 5 mutations in act1 and rp51A are complementary to these invariant U5 nucleotides, suggesting that the abnormal cleavage events may be favored by interactions with both U6 and U5.

In summary, three of the five spliceosomal snRNAs (U1, U5, and U6) can participate in the identification and use of the 5' splice site. On the basis of kinetic analysis (*9,10*) and the mutually exclusive base-pairing interactions (Fig. 1 A compared to C and D), it seems likely that the U1 and U6 interactions with the 5' splice site are sequential (*10,14*), while, in contrast, the U5 and U6 interactions may occur simultaneously.

RNA-RNA interactions with U6 may juxtapose the branch point and the 5' and 3' splice sites. A model for the spliceosomal active site has been proposed based on genetic suppression studies in which a previously undescribed pairing interaction between U6 and U2 (helix I) was identified (*18*). Because formation of this helix is mutually exclusive with the pairing of U6 to U4, which is thought to be the negatively regulated form of U6 (*16*), this rearrangement could correspond to the catalytic activation of the spliceosome (*18*). The further significance of helix I derived from its location immediately upstream of the branch-point recognition sequence in U2 and immediately downstream of the highly conserved (*16,17*) and essential (*19,20*) nucleotides of the ACAGAG sequence in U6. However, this model left unspecified how the 5' splice site could be delivered to the presumptive catalytic core. The newly demonstrated base-pairing interaction between the invariant ACAGAG sequence and the 5' splice site (Fig. 7) can serve the critical function of aligning the 5' splice site of the substrate with the branch point nucleotide, which is the attacking nucleophile in the first chemical step (*21*).

The configuration shown in Fig. 7 can also accommodate RNA-RNA interactions that occur in the second catalytic step of splicing. Extension of the short 5' splice site-U6 helix would result in the apposition of the last three residues of the ACAGAG sequence (G50 to G52) with the first three nucleotides of the intron (/GUA). Consistent with this alignment, recent mammalian crosslinking experiments (*22*) have provided evidence that the last A of the ACAGAG sequence (A51 in yeast) is in close proximity to intron position 2 (/GU). This A-U crosslink is induced in the lariat intermediate but not in the pre-mRNA. Furthermore, as described above, genetic suppression studies have demonstrated interactions between G1 and G303, the first and last nucleotides of the actin intron, during the second catalytic step (*23*). These studies imply that the first and last positions of the intron and the last two nucleotides of the ACAGAG sequence are in proximity during the second step of splicing. This is consistent with the previous demonstration that mutations of residues 51 and 52 of ACAGAG inhibit the second step of splicing of a wild-type actin intron in vitro (*20*).

We have now demonstrated suppression of mutations in the last two positions of the actin intron (AG/) by mutation of G52 in the ACAGAG sequence. The U6-G52U mutation does not exhibit allele specificity; that is, it can suppress the second step block in all G303 mutations as well as one A302 mutation. The more restricted allele specificity of the G1-G303 suppression spectrum probably indicates a direct, but non-Watson-Crick, interaction between these two nucleotides (*23*). Whether the suppression of 3' splice site

mutations by U6-G52U also reflects a direct, albeit complex, set of RNA-RNA interactions bridging G52 to both A302 and G303 cannot be determined by our present experiments. Nonetheless, this interpretation is at least consistent with the genetic and biochemical data linking actin intron G1 to G303 (*23*) and U6-A51 to actin intron U2 (*22*), respectively.

In summary, the 5' and 3' splice sites can be modeled as juxtaposed with one another and with the ACAGAG region of U6 by a set of RNA-RNA interactions (Fig. 7) which appear to comprise Watson-Crick as well as noncanonical pairings (*23,46*). Since the interactions described above encompass both the first and second chemical steps, this model suggests that the catalytic sites for the two steps might be substantially overlapping (*18*), although certainly not identical (*21,47*).

REFERENCES AND NOTES

1. M. R. Green, *Annu. Rev. Cell Biol.* **7**, 559 (1991).

2. C. Guthrie, *Science* **253**, 157 (1991); M.J. Moore, C.C. Query, P.A. Sharp, in *The RNA World* , R. Gesteland and J. Atkins, Eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1993) pp. 303-58.

3. Y. Zhuang and A. M. Weiner, *Cell* **46**, 827 (1986).

4. P. G. Siliciano and C. Guthrie, *Genes Dev.* **2**, 1258 (1988); B. Seraphin, L. Kretzner, M. Rosbash, *EMBO* **7**, 2533 (1988).

5. R. Parker, P.G. Siliciano, C. Guthrie, *Cell* **49**, 220 (1987); Y. Zhuang and A. M. Weiner, *Genes Dev.* **3**, 1545 (1989); J. Wu and J. L. Manley, *ibid.* **3**, 1553 (1989).

6. P. Legrain, B. Seraphin, M. Rosbash, *Mol. Cell. Biol.* **8**, 3755 (1988); B. Seraphin and M. Rosbash, *Cell* **59**, 349 (1989); S. Michaud and R. Reed, *Genes Dev.* **5**, 2534 (1991); S. Jamison, A. Crow, M. Garcia-Blanco, *Mol. Cell. Biol.* **12**, 4279 (1992).

7. A.J. Newman and C. Norman, *Cell* **65**, 115 (1991).

8. A.J. Newman and C. Norman, *ibid.* **68**, 743 (1992).

9. J.R. Wyatt, E.J. Sontheimer, J.A. Steitz, *Genes Dev.* **6**, 2542 (1992).

10. D.A. Wassarman and J.A. Steitz, *Science* **257**, 1918 (1992).

11. J.L. Woolford, *Yeast* **5**, 439 (1989); B. C. Rymond and M. Rosbash, in *Molecular and Cellular Biology of the Yeast, Saccharomyces: Gene expression* J. R. Broach, J. R. Pringle, E. W. Jones, Eds. (Cold Spring Harbor Press, Cold Spring Harbor, NY, 1992), pp. 143-192.

12. A slash mark (/) indicates a site of cleavage. If the (/) preceeds the sequence it represents the site of 5' cleavage and if the (/) follows the sequence it represents the site of 3' cleavage.

13. R. Parker and C. Guthrie, *Cell* **41**, 107 (1985); A. Jacquier, J. R. Rodriguez, M. Rosbash, *ibid.* **43**, 423 (1985); L. A. Fouser and J. D. Friesen, *ibid.* **45**, 81 (1986).

14. H. Sawa and J. Abelson, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 11269 (1992).

15. Because of limitations in the resolution of the cross-linked site in the intron, the data of Wassarman and Steitz (*10*) are also consistent with the model proposed by Sawa and Abelson (*14*). Alternatively, both pairing schemes may be correct but reflect interactions at different steps in the pathway; this interpretation is

supported because U6 has been found crosslinked to the pre-mRNA in one case (*10*) and to the lariat intermediate in the other (*14*).

16. C. Guthrie and R. B. Patterson, *Annu. Rev. Genet.* **23**, 387 (1988).

17. C. Guthrie, S. Mian and H. Roiha, unpublished data.

18. H.D. Madhani and C. Guthrie, *Cell* **71**, 803 (1992).

19. H. D. Madhani, R. Bordonné, C. Guthrie, *Genes Dev.* **4**, 2264 (1990).

20. P. Fabrizio and J. Abelson, *Science* **250**, 404 (1990).

21. M. J. Moore and P. A. Sharp, *Nature* , **365**, 364 (1993).

22. E.J. Sontheimer and J.A. Steitz, *Science,* in press.

23. R. Parker and P.G. Siliciano, *Nature* **361**, 660 (1993).

24. U6 point mutations were generated by polymerase chain reaction (PCR) amplification with primers that contained the desired nucleotide changes (*18*). The mutant U6 genes were cloned into the pSX6 low copy centromere containing vector (*18*).

25. Intron point mutations were generated by PCR amplification with primers that contained the desired nucleotide change (*28*). The

ACT1-CUP1 intron fusions under the control of the glucose-6-phosphate-dehydrogenase (GPD) promoter (*28*) were cloned into the high copy (2μ) RS423 or RS426 vectors [R.S. Sikorski and P. Heiter, *Genetics* **122**, 19 (1989)].

26. Specific intron or snRNA mutations are represented by a number flanked by letters. The number represents the position of the mutation in the RNA. The letter preceeding the number represents the wild-type sequence and the letter following defines the mutation.

27. Cleavage at the aberrant site normally results in the accumulation of lariat intermediate, because a G is required at intron position 1 in the second catalytic step 2 [B. Seraphin and M. Rosbash, *Cell* **63**, 619 (1990)]. We have changed the U at position -5 in actin (the aberrant cleavage site) to G. This allows the aberrant lariat intermediate to proceed through the second catalytic step, and a mRNA five nucleotides shorter than the normal mRNA is produced.

28. C.F. Lesser and C. Guthrie, *Genetics* **133**, 851 (1993).

29. C.F. Lesser and C. Guthrie, unpublished data.

30. Although the mutations introduced at U6 positions 45 and 46 should create two new contiguous base pairs in model I, they only negligibly strengthen the base-pairing interaction proposed in model

II, in that they would introduce only a single noncontiguous A·U or G·U base pair.

31. Primer extension products were quantitated by means of a phosphor-imaging system (Molecular Dynamics), and the data were analyzed using the ImageQuant software. The numerical results are based on analysis of three independent transformants. The degree of suppression was determined by comparing the change in the ratio of mRNA generated by cleavage at the normal site relative to the aberrant site in one strain relative to another.

32. A comparison of the sequence of 35 *S. cerevisiae* introns reveals no sequence conservation at intron positions 7 and 8 (D.N. Frank and C. Guthrie, unpublished data).

33. None of these U6 point mutants conferred a detectable growth phenotype when assayed at 18°, 25°, 30°, or 37°C, nor did they influence splicing of a wild-type intron reporter as assayed by copper resistance.

34. That the U6-C43U mutant does not increase aberrant cleavage activated by the intron G5A mutation may be due to the greater stability provided by a G·C base pair in wild-type and U6-C43G-intron-G5C interactions.

35. Mutations in this region decreased the splicing efficiency of introns mutated at position 5, but did not greatly alter the relative amount of cleavage at the aberrant compared to the normal site.

36. Primer extension analyses confirmed that cleavage was increased at the normal site and decreased at the aberrant site.

37. Y-T. Yu, P. Maroney, T. Nilsen, *Cell*, in press.

38. A. J. Newman, R.J. Lin, S. Cheng, J. Abelson, *Cell* **42**, 335 (1985); U. Vijayraghavan *et al.*, *EMBO J.* **5**, 1683 (1986).

39. L. A. Fouser and J. D. Friesen, *Cell* **45**, 81 (1986).

40. All actin 3' splice site mutations were also mutated at position 305 of the pre-mRNA to prevent activation of an aberrant cleavage site (*23*).

41. The intron position 1 and 303 (or both) mutant ACT-CUP1 reporter constructs were provided by P. Siliciano. The construction of these plasmids has been described (*23*). These ACT1-CUP1 fusions contain about 50 additional nucleotides of endogenous actin exon 2 as compared to the ACT1-CUP1 reporters used earlier in this work.

42. Primer extension analyses have shown that when a wild-type strain is transformed with an ACT1-CUP1 reporter which has been

mutated at position 303 (G303C,G305C), cleavage occurs very weakly at the mutant 3' splice site and at an aberrant 3' splice site (UG/) located upstream from the normal site (*23*). U6-G52U increases cleavage at both the mutated 3' splice site (G303C) and also at the aberrant 3' cleavage site (UG/) located 5 nucleotides upstream from the normal site. In contrast, the intron G1A mutation only increases cleavage at the normal site. The ability of U6-52U to increase cleavage at both the mutated and aberrant sites is consistent with the observation that this mutation does not exhibit allele-specific suppression.

43. Primer extension analyses conducted on RNA isolated from U6-ACAGAG mutant strains transformed with the G1A ACT1-CUP1 reporter confirmed that the G50C and G52U strains decrease the efficiency of the first catalytic step when the intron is mutated at position 1 of the 5' splice site.

44. S. Kandel and B. Seraphin, *Science,* in press. U6 mutants were expressed on high copy plasmids to detect efficient suppression of intron position 5 mutations on low copy reporters. Since our reporter is high copy and U6 construct is low copy, it is not surprising that we cannot detect suppression in heterozygotes.

45. Sawa and Abelson (*14*) have pointed to the possibility that mutation of intron position 5 could lead to incorrect positioning of the catalytic center as a result of disruption of a U6-intron interaction.

46. The 4-thioU crosslink between intron position 2 and the last A of the ACAGAG sequence may well be non-Watson Crick, E. Sontheimer and J.A. Steitz, *Science*, in press.

47. J.A. Steitz and T. A. Steitz, *Proc. Natl. Acad. Sci. U.S.A.* **90**, 6498 (1993).

48. S.L. Johnson, thesis, University of Washington, Seattle (1991).

49. Total RNA was analyzed by a primer extension method using a ^{32}P-end-labeled oligonucleotide complementary to CUP1 which is able to detect pre-mRNA, lariat intermediate and mature mRNA (*28*). A labeled oligonucleotide complementary to U1 was also included in the reaction as an internal control for the amount of RNA in each reaction (*28*). The products were analyzed by electrophoresis through a 6% denaturing acrylamide gel followed by autoradiography.

50. The approximate size of this band is consistent with the hypothesis that the new G at position 8 is now used as an aberrant 5' splice site; further experiments are needed to be done to test this hypothesis.

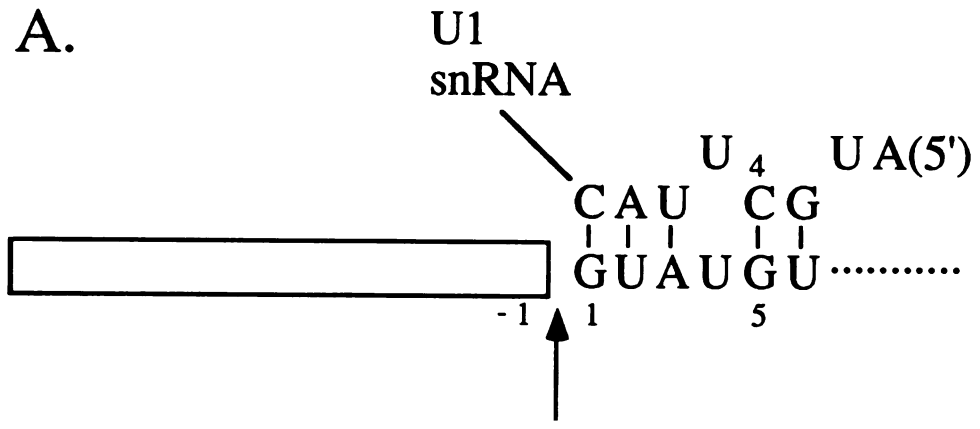
51. T. P. Hausner and A. M. Weiner, *Genes Dev.* **4**, 2146 (1990); B. Datta and A. M. Weiner, *Nature* **352**, 821 (1991); J. A. Wu and J. L. Manley, *ibid.* **352**, 818 (1991).

52. T. Wolff and A. Bindereiff, *Genes Dev.* **7**, 1377 (1993).

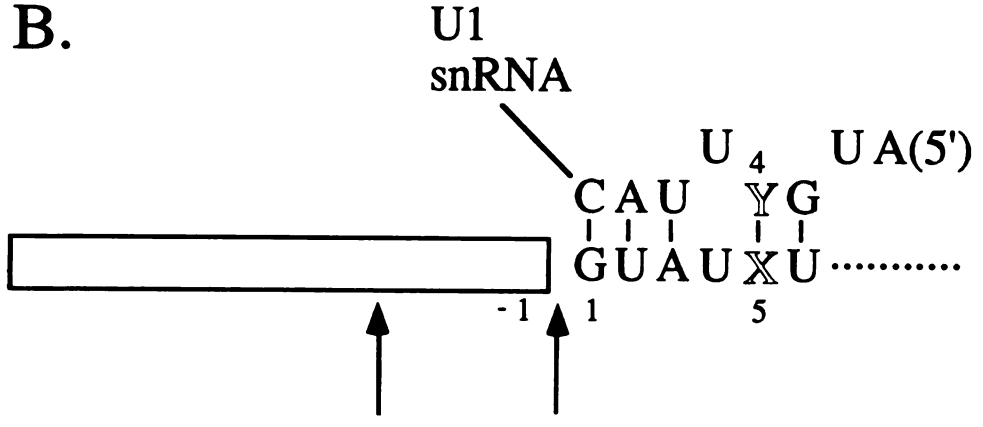
53. We thank H. Sawa and J. Abelson for sharing information prior to publication and for their continued contributions to this work; H.D. Madhani, J.P. Staley, S. Burgess, P. Siliciano, and A. Murray for helpful discussions; H.D. Madhani, J.P. Staley, R. Parker, E. Blackburn, P. Walter, J. Abelson, and E. Sontheimer for criticism of the manuscript; and E.J. Sontheimer, J.A. Steitz, S. Kandel-Lewis, B. Seraphin, Y-T. Yu, P. Maroney, and T. Nilsen for communication of unpublished results. This work was supported by grant GM21119 from the National Institutes of Health.

Figure 1. Base-pairing interactions between the 5' splice site and U1 and U6 snRNAs. Shown is the complementarity between the *S. cerevisiae* 5' splice site consensus sequence and either U1 (A,B) or U6 (C,D). The open rectangle represents exon 1; the arrow between positions 1 and -1 indicates the normal cleavage site. **(A)** The 5' end of U1 is complementary to and base pairs with the 5' splice site. **(B)** Disruption of this interaction by intron position 5 mutations (X) results in cleavage at an aberrant splice site upstream of the 5' splice site as represented by the upstream arrow. Restoration of complementarity by the introduction of point mutations in U1 at position 4 (Y), does not prevented aberrant cleavage (see text for references). **(C)** Model I (*10*) proposes that the invariant ACA sequence of U6 (nt. 47 to 49 in yeast) base pairs with positions 4 to 6 of the 5' splice site. **(D)** Model II (*14*) proposes that an upstream ACA sequence of U6 (ntes 42-44 in yeast) base pairs with positions 4-6 of the 5' splice site.

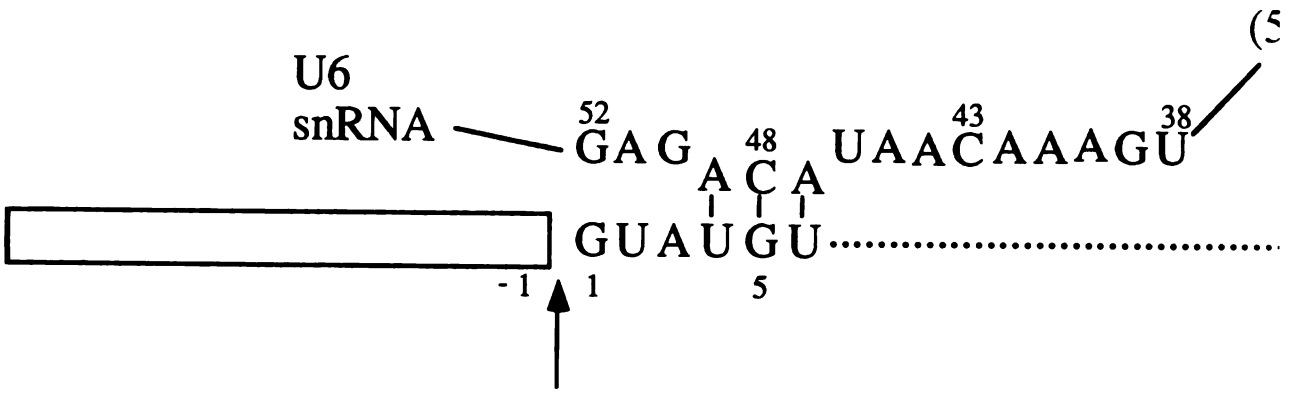
A.


U1
snRNA
U 4
UA(5')
CAU CG
GUAUGU
-1
1
5

B.


U1
snRNA
U 4
UA(5')
CAU YG
GUAUXU
-1
1
5

C. MODEL I



D. MODEL II

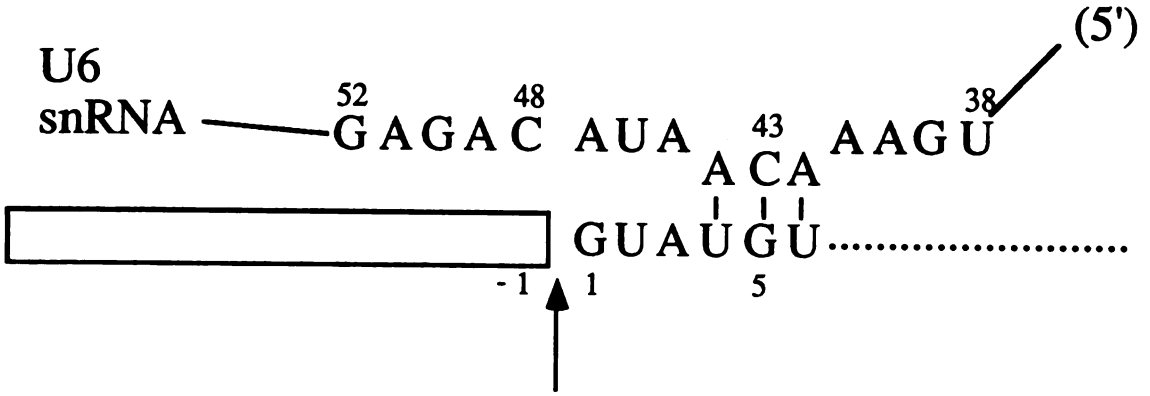


Figure 2. General strategy for assaying effects of U6 point mutations on intron mutations. **(A)** Shown is the yeast strain YCL47 which contains a deletion of both the chromosomal copies of U6 and CUP1. Low copy (cen) plasmids carrying U6 point mutations (*)(*24*) were introduced into this strain by plasmid shuffle methods (*18*) and then ACT1-CUP1 reporters on high copy (2 micron) plasmids (*25*) were introduced into the strain by standard yeast transformation protocols and copper-resistance levels were determined (*28*). **(B)** The wild-type reporter plasmid measures cleavage at the normal (NL) cleavage site in the absence of any 5' splice site mutations. The white boxes represent ACT1 sequences which flank the intron; the filled-in boxes represent CUP1 sequence (*28*). Depicted is a pair of reporter constructs, containing intron position 5 mutations (*), whose copper-resistance reflect cleavage at the normal (NL) **(C)** or aberrant (AB)**(D)** cleavage sites (*27*).

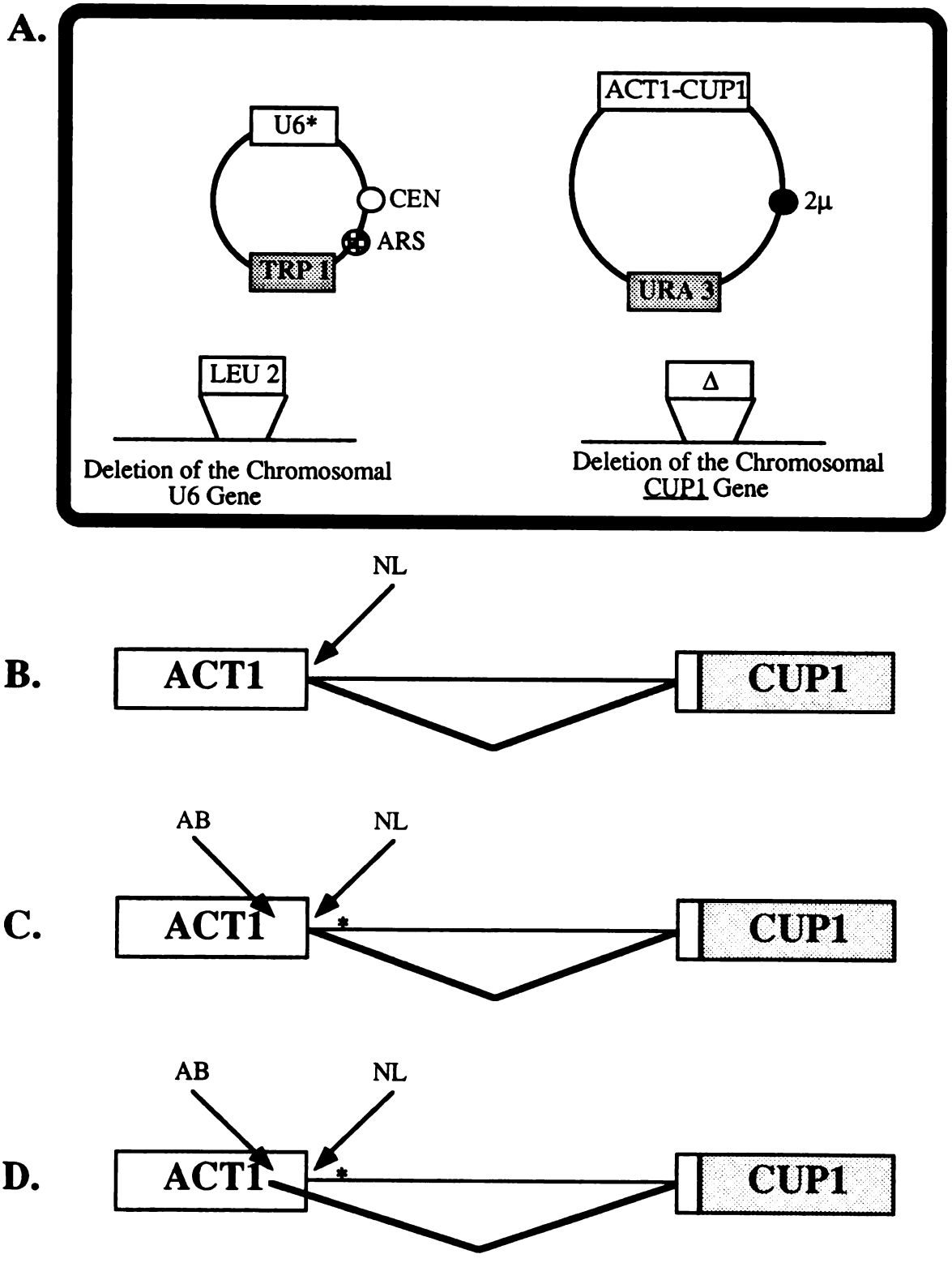
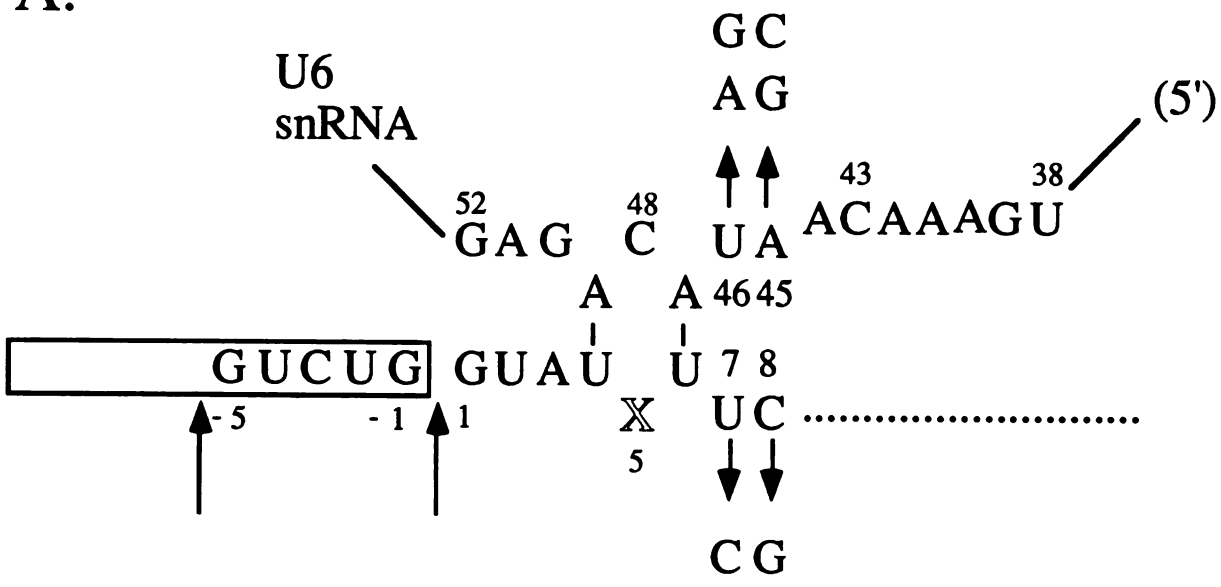
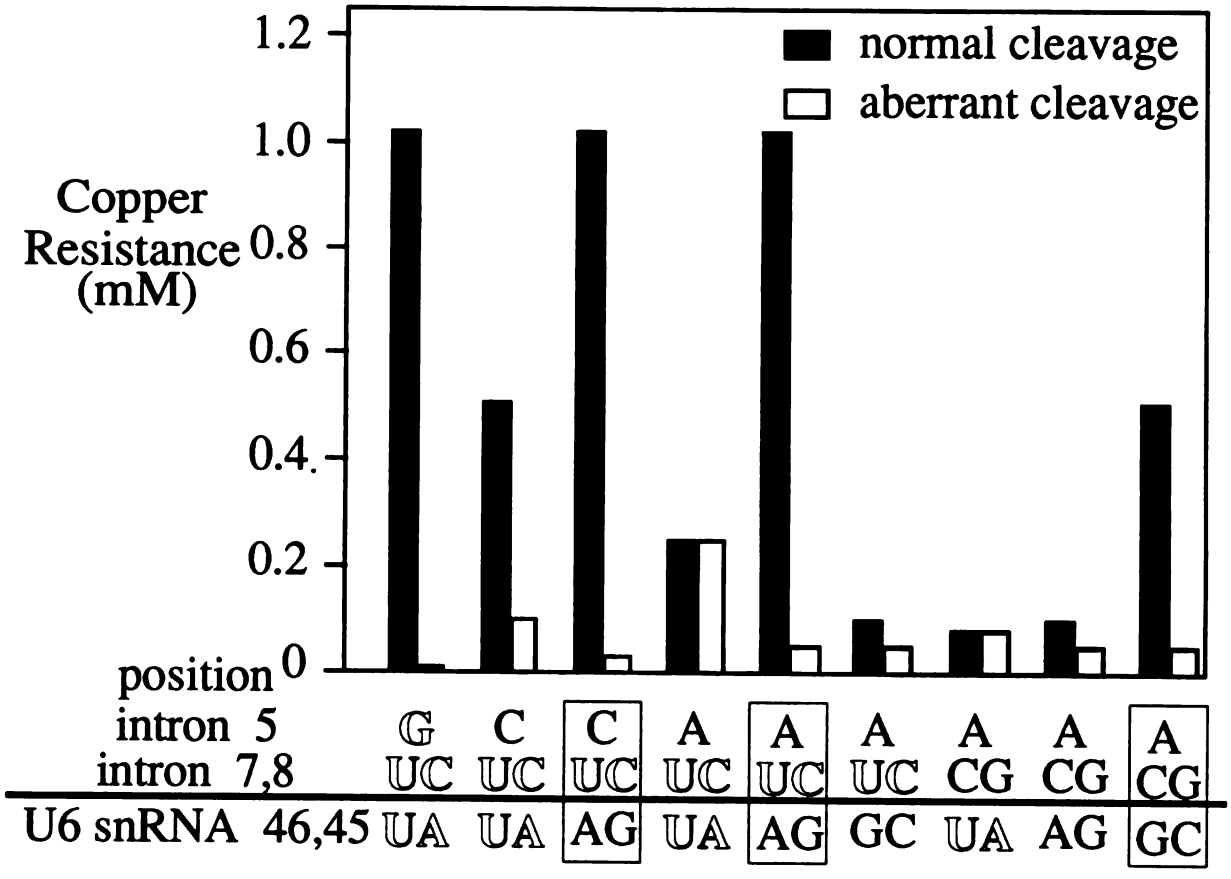

A.
U6*
CEN
ARS
TRP 1
ACT1-CUP1
2μ
URA 3
LEU 2
Δ
Deletion of the Chromosomal U6 Gene
Deletion of the Chromosomal CUP1 Gene
B.
NL
ACT1
CUP1
C.
AB
NL
ACT1
CUP1
D.
AB
NL
ACT1
CUP1

Figure 3. A test of the base-pairing interaction proposed in Model I (*10*). **(A)** Disruption of the potential U6-5' splice site base-pairing interaction proposed in model I by the introduction of intron position 5 mutations. Mutations at position 5 (X), result in activation of aberrant cleavage at position -5 as shown by the upstream arrow. Mutations were introduced at U6 positions 45 and 46 and actin intron positions 7 and 8 to test the prediction that increasing the base-pairing region between these two RNAs should decrease activation of aberrant cleavage. **(B)** The copper resistance of U6-wild-type and U6-position 45,46 mutant strains transformed with ACT1-CUP1 reporters. The first two bars of the table reflect copper resistance of a U6-wild-type strain transformed with a wild-type intron ACT1-CUP1 reporter. At the bottom of the graph is indicated the specific sequence of the intron reporter and U6; the boxed combinations indicate when complementarity between U6 and the 5' splice site should be restored. The bold underlined nucleotides represent wild-type sequence. **(C)** Primer extension analysis of RNA isolated from U6-wild-type or U6-45,46 mutant strains carrying intron position 5,7,8 mutant ACT1-CUP1 reporters. Each lane is labeled with the specific genotype of the strain (U6 allele and intron reporter construct) from which the RNA was isolated. The products of the primer extension assays (*49*) are indicated. RNA was isolated from strains carrying a reporter plasmid which generates an in-frame mRNA when cleavage is at the aberrant site. Since out-of-frame mRNAs are significantly less stable than in-frame mRNAs, the ratio of the levels of aberrant mRNA relative to normal mRNA is skewed in favor of the aberrant mRNA (*28*). The * marks a new band

which appears when the intron U7C,C8G mutations are introduced (*50*).

A.


U6
snRNA
(5')
52
48
43
38
GAG C UA ACAAAGU
A A 46 45
GUCUG GUAU U 7 8
-5 -1 1 X UC
5
CG


normal cleavage
aberrant cleavage
Copper Resistance (mM)
1.2
1.0
0.8
0.6
0.4
0.2
0
position
intron 5
intron 7,8
U6 snRNA 46,45
G C C A A A A A A
UC UC UC UC UC UC CG CG CG
UA UA AG UA AG GC UA AG GC

C.

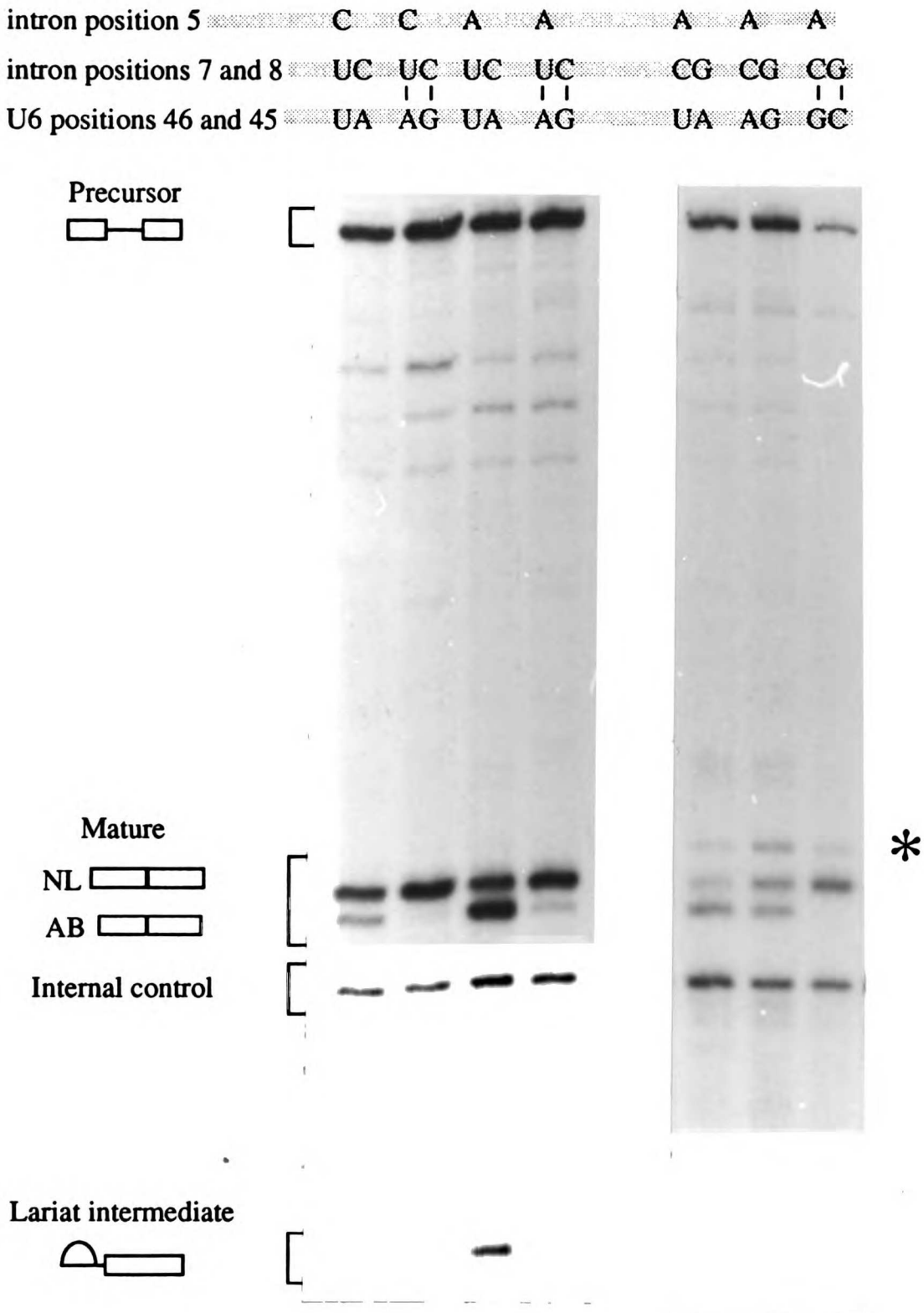

intron position 5 C C A A A A A
intron positions 7 and 8 UC UC UC UC CG CG CG
U6 positions 46 and 45 UA AG UA AG UA AG GC
Precursor
Mature
NL
AB
Internal control
Lariat intermediate
*

Figure 4. A test of the base-pairing interaction proposed in model II (*14*). **(A)** The extended base-pairing interaction that can be drawn between the actin intron and U6 (model II). **(B)** The pairing in (A) may be favored when mutations are introduced at intron position 5 (G5C), disrupting the normal base-pairing interaction proposed in model I (Fig. 1C). This alternative base-pairing interaction could activate aberrant cleavage as indicated by the arrow at -5. **(C)** Increasing the complementarity between the G5C intron and U6 by the introduction of the C43G mutation, increased cleavage at the aberrant site and decreased cleavage at the normal site (Fig. 5) as depicted by the thickness of the arrows demarcating the two alternative cleavage sites.

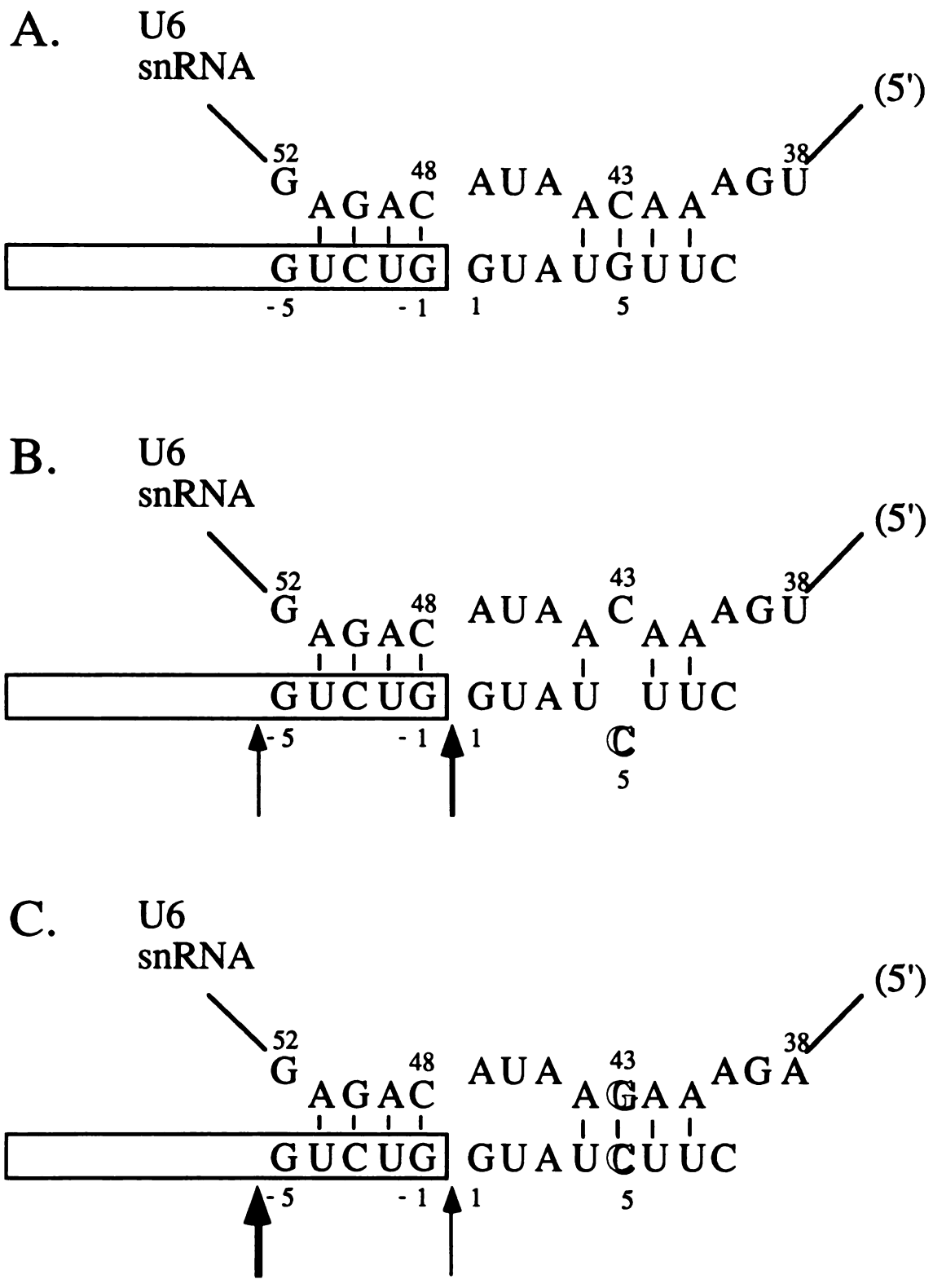

A.
U6
snRNA
(5')
52
48
43
38
G AGAC AUA ACAA AGU
GUCUG GUAUGUUC
-5
-1
1
5
B.
U6
snRNA
(5')
52
48
43
38
G AGAC AUA A C AA AGU
GUCUG GUAU UUC
C
-5
-1
1
5
C.
U6
snRNA
(5')
52
48
43
38
G AGAC AUA AGAA AGA
GUCUG GUAUCUUC
-5
-1
1
5

Figure 5. Primer extension analysis of RNA isolated from U6 position 43 mutant strains transformed with G5A and G5C reporters. Total RNA was isolated from either a wild-type or position 43 mutant U6 strain transformed with the G5A or G5C reporter which generates an in-frame mRNA when cleavage is at the aberrant site. Each lane is labeled with the specific U6 strain and the intron reporter from which the RNA was isolated and the products of the primer extension assays are indicated on the side of the figure (*49*).

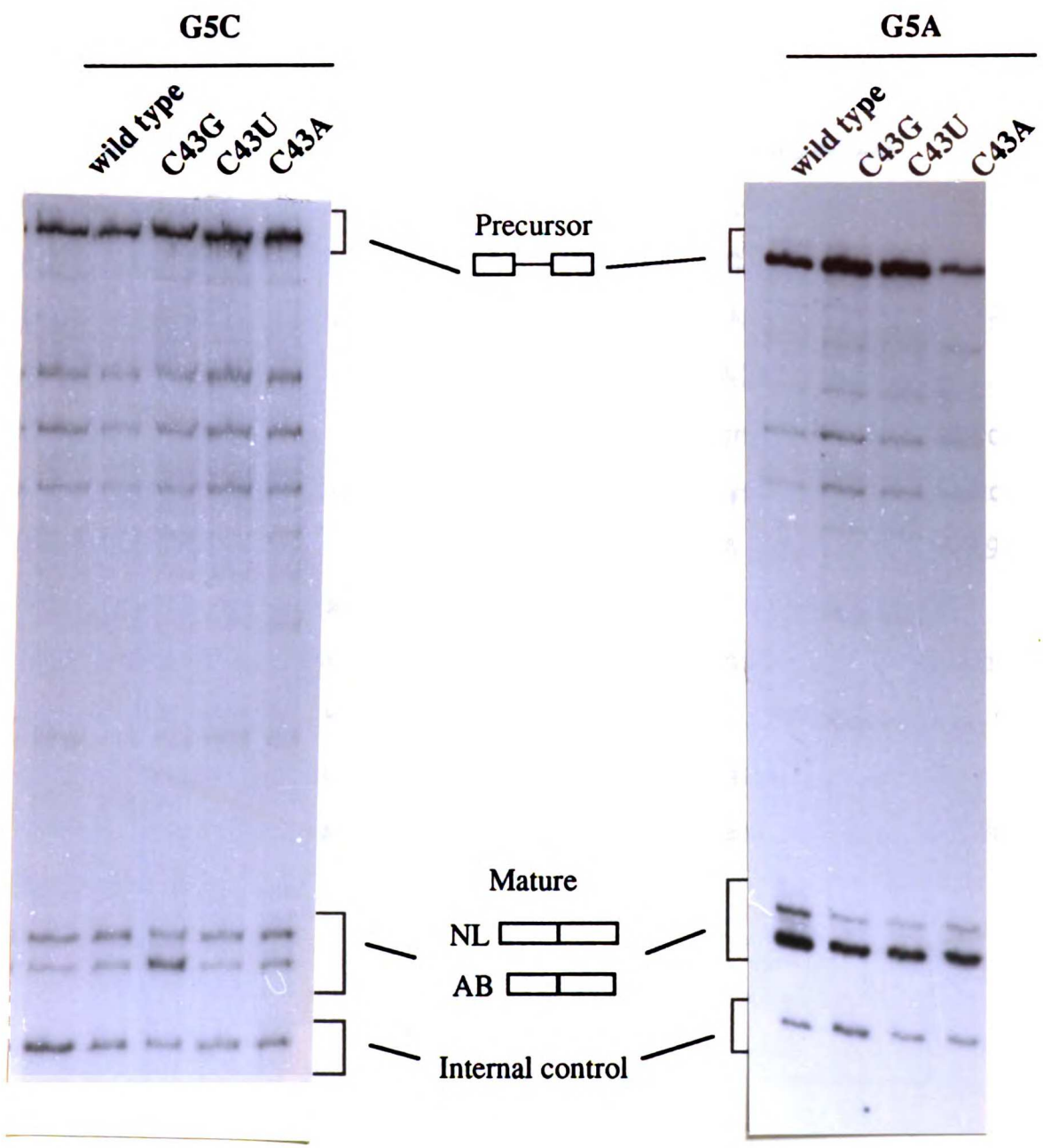
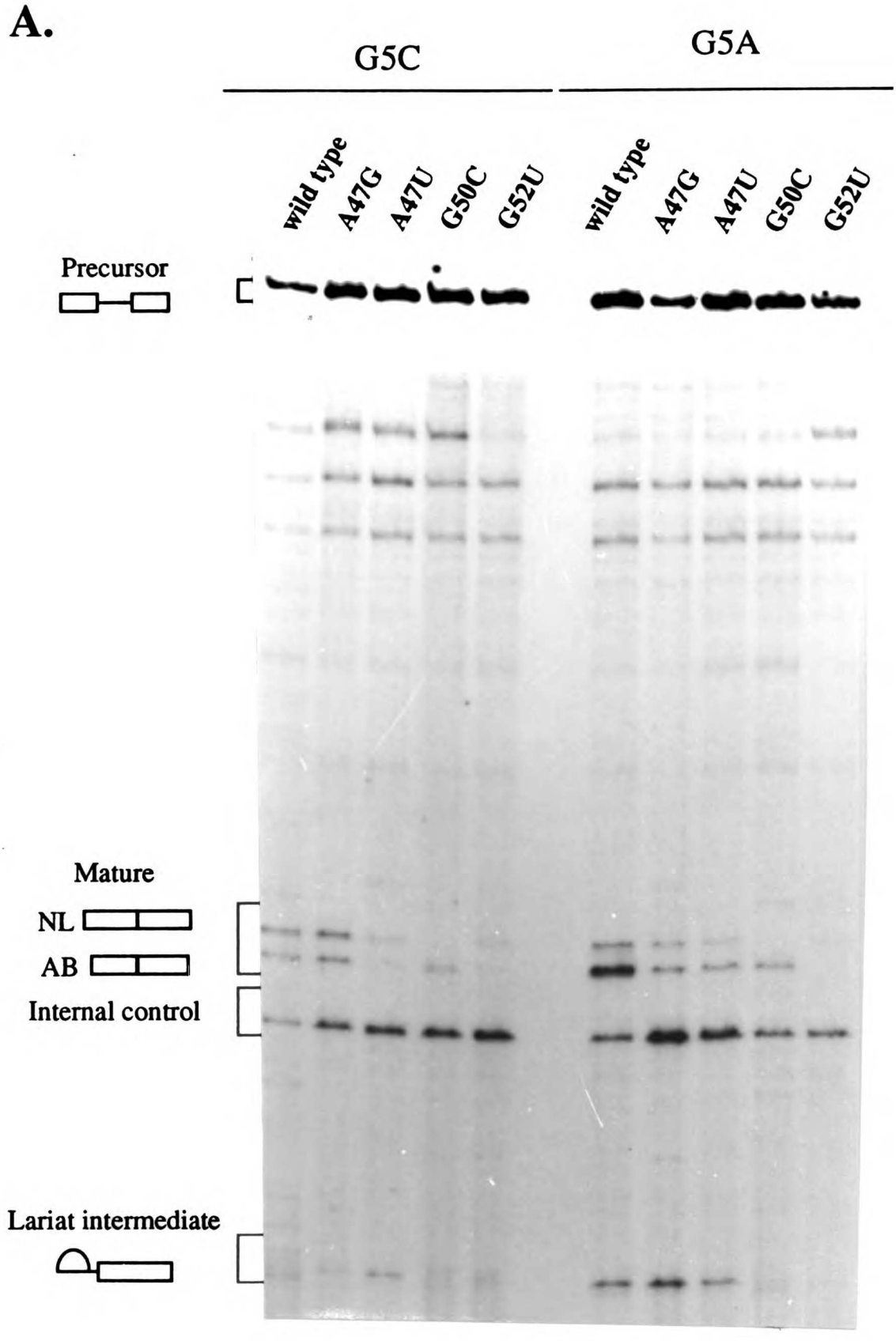

G5C
G5A
wild type
C43G
C43U
C43A
wild type
C43G
C43U
C43A
Precursor
Mature
NL
AB
Internal control

Figure 6. Primer extension analyses of the RNA isolated from U6-ACAGAG mutant strains transformed with 5' and 3' splice site mutant reporters. **(A)** Splicing of the intron G5C and G5A reporters for aberrant cleavage by strains mutated in the U6 ACAGAG sequence. **(B)** Splicing of the intron G303C reporter (*23,40,41*) by strains mutated in the U6 ACAGAG sequence. **(C)** Splicing of the intron G1A-G303C reporter (*23,40,41*). Total RNA was isolated from strains containing either U6-wild-type or a U6-ACAGAG mutant transformed with the ACT1-CUP1 reporter designated at the top of each panel. Each lane is labeled with the genotype with respect to the U6 of the specific strain from which the RNA was isolated (*49*). The products of the primer extension assays are indicated on the side of the figure. On longer exposures it is possible to detect both normal and aberrant mRNAs in the G52U lane in (A) in addition to the mRNAs resulting from cleavage at both the mutated 3' splice site (G303C) and the aberrant site located five nucleotides upstream in the intron (*42*) in (B).


A.
G5C
G5A
wild type
A47G
A47U
G50C
G52U
wild type
A47G
A47U
G50C
G52U
Precursor
Mature
NL
AB
Internal control
Lariat intermediate

B.

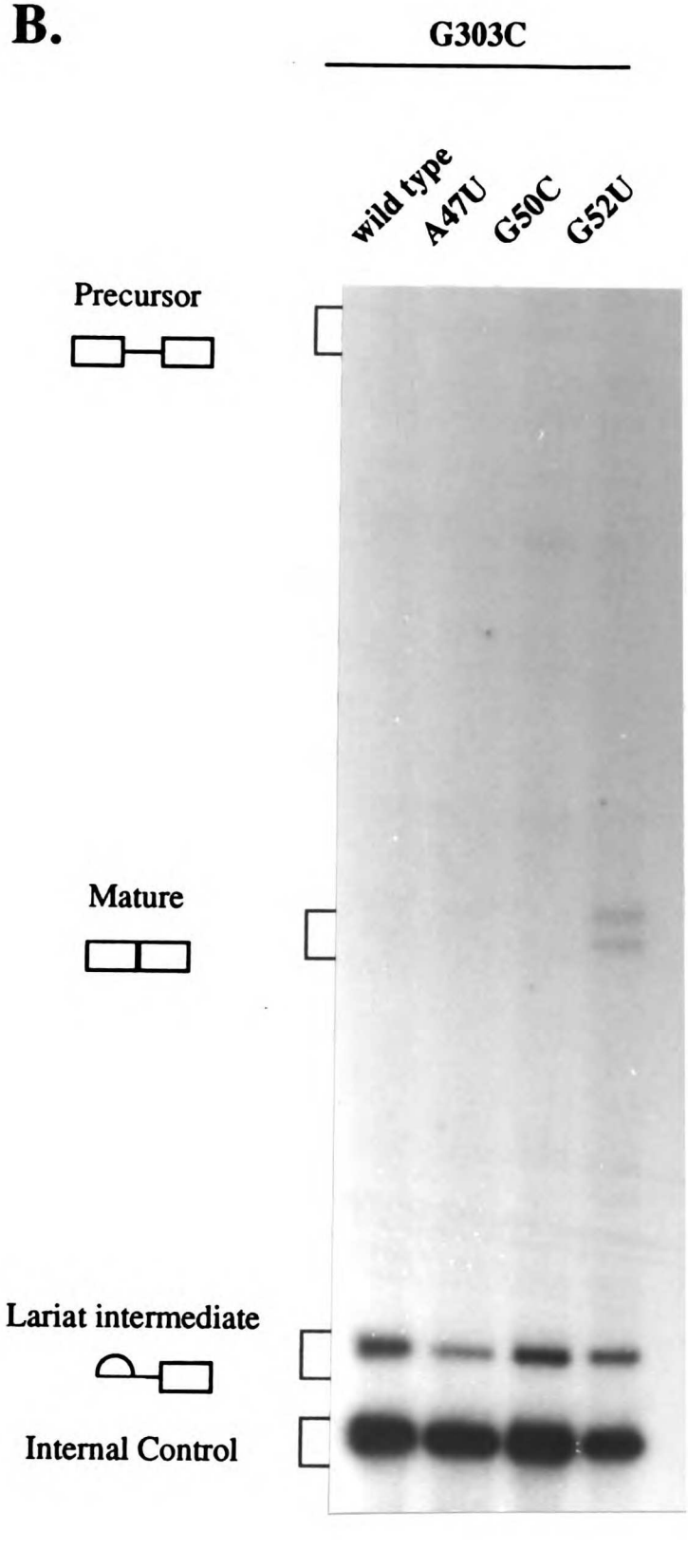

G303C
wild type
A47U
G50C
G52U
Precursor
Mature
Lariat intermediate
Internal Control

Figure 7. U6 base-pairing interactions juxtapose the branch point A with the 5' and 3' splice sites. The pre-mRNA is shaded and the 5' and 3' splice site nucleotides as well as the branch point A are drawn in a larger font and are numbered according to their positions in the yeast actin gene. E1 and E2 represent exon 1 and exon 2, respectively, and the U2 and U6 sequences are labeled. The proposed base-pairing interaction between U6 positions 47-49 and intron positions 4-6, described in this manuscript (Model I), is shown along with four other previously identified helices, indicated in smaller font or line drawings. The interactions between U6 and the intron can act to juxtapose the splice sites in close proximity to the putative catalytic residues of the spliceosome. The four previously identified RNA·RNA helices in the spliceosome are: (*i*) the branch point recognition helix (*5*) (*ii*) the U2·U6 helix I (*18*) (*iii*) the U2·U6 helix II (*51*) and (*iv*) the intramolecular U6 helix (*16,17, 52*).

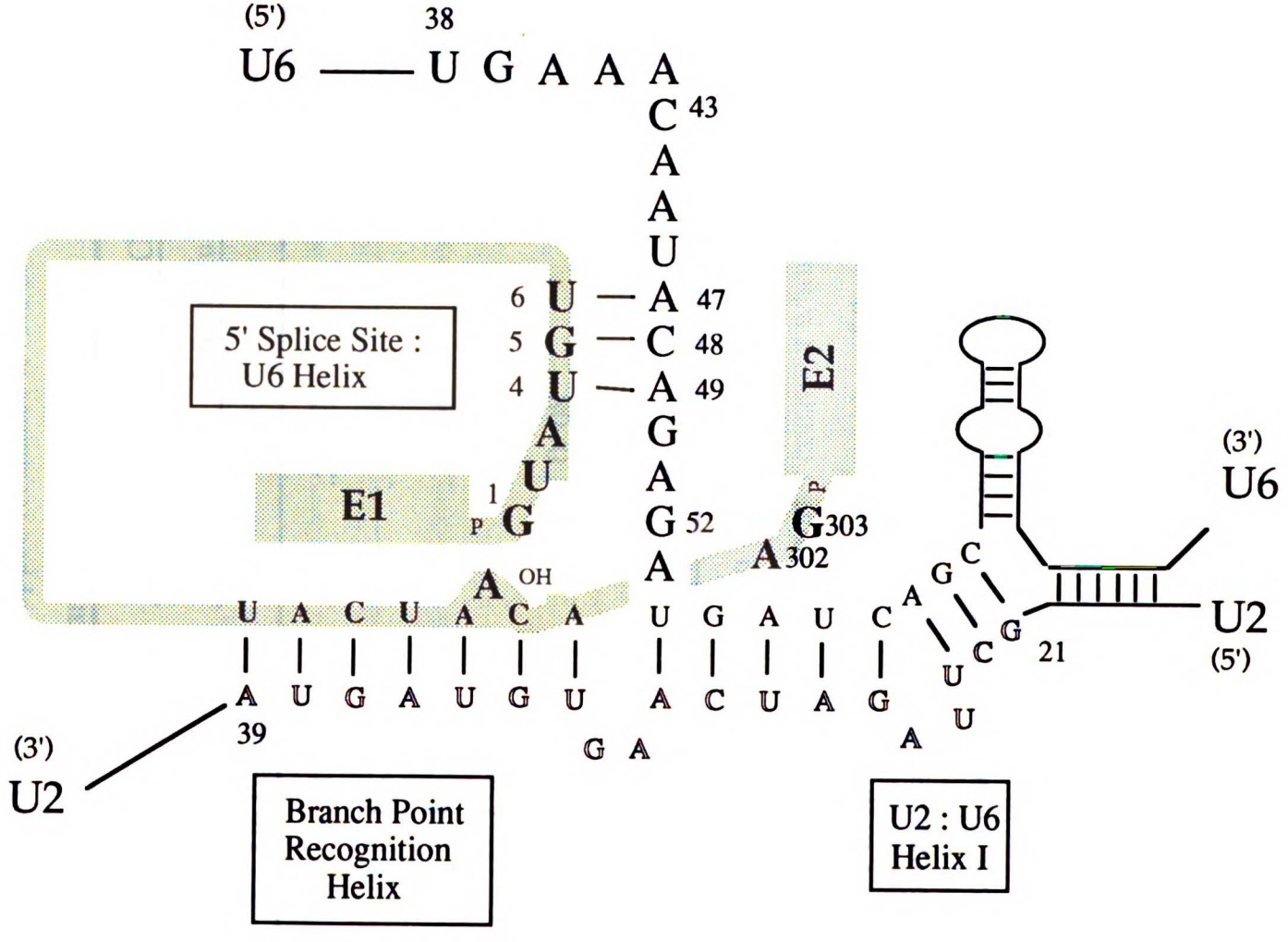

(5')
38
U6
U G A A A
C 43
A
A
U
6 U — A 47
5 G — C 48
4 U — A 49
A
U
G
A
G 52
A
5' Splice Site :
U6 Helix
E1
E2
1
P
G
P
G303
A302
A OH
U A C U A C A
U G A U C
A G C
G
C
U
C G
21
U2
(5')
(3')
U6
A U G A U G U A C U A G
A U
39
G A
(3')
U2
Branch Point
Recognition
Helix
U2 : U6
Helix I

Table 1 Summary of copper-resistance levels of U6 strains transformed with either normal or aberrant reporters.

	COPPER-RESISTANCE LEVELS (mM)*	
	Intron Mutation	
U6 snRNA	G5C	G5A
wild type	.50/.10	.25/.25
C43G	.10/.13	.10/.08
C43U	.25/.08	.18/.10
C43A	.50/.10	.25/.18
G39C	1.0/.08	.50/.08
G39A	.50/.10	.50/.25
G39U	.50/.10	.50/.25

* The top value of the ratio represents the level of copper-resistance of a strain transformed with a reporter construct whose growth reflects cleavage at the normal site, while the bottom value of the ratio represents the level of copper resistance of the same strain transformed with a reporter construct whose growth reflects cleavage at the aberrant site (23).

Table 2. Effects of U6-ACAGAG mutations on 5' and 3' splice site mutations in the actin intron.

```
1                                       302 303 305*
|                                         \  |   |
G/GUAUGU------------------UACUAAC---------UAG/AG
```

Intron Mutation ‡	β-Galactosidase Levels † — U6 snRNA Mutant: wild type	A47U	G50C	G52U
G1A	-	-	-	-
G303C	+	-	-	++
G303U	++	++	+	+++
G303A	+/++	+	-	+++
A302C	++	++	+	+++
A302U	++	++	-	++
G1A/G303C	+++	++	-	+
G1A/G303U	+/-	-	-	-
G1A/G303A	+	-	-	-

* All 3' splice site mutations were analyzed in a G305C background to avoid activation of an aberrant cleavage site at actin exon positions 305.

† β-galactosidase levels were determined by triplicate overlay assays (45)

‡ The U6 snRNA mutants do not affect splicing of a wild intron actin-LacZ fusion.

Chapter 6

U5 snRNA mutations influence cleavage site choice

As discussed in Chapter 1, the 5' splice site consensus sequence, the last three nucleotides of the exon and the first six of the intron, are complementary to the 5' end of U1 snRNA. Furthermore, U1 and the 5' splice site have been demonstrated to base-pair and this interaction, in *S.cerevisiae* , is sufficient to commit the pre-mRNA to the splicing pathway. In addition, as discussed in Chapter 5, I have shown that the last three nucleotides of the 5' splice site potentially base pair with an invariant region of U6 snRNA. Notably, this short helix plays an important role in defining the exact location of cleavage at the 5' exon-intron boundary. The U1·5' splice site and U6·5' splice site base-pairing interactions are mutually exclusive. Recent experiments conducted in our laboratory and others indicate that under special conditions, U5 snRNA can influence 5' splice site selection. This data, reviewed below, suggests that one role of U5 snRNA might be to maintain 5' splice site definition after U1 has left and before U6 has bound.

Genetic interactions between U5 snRNA and the 5' splice site. The first indication that snRNAs in addition to U1 might interact with the 5' splice site was the result of a genetic screen to identify suppressors of the cyh2 intron position 1 G1A mutation. Newman and Norman (1991) discovered unexpectedly that a mutation in the invariant loop of U5 snRNA could activate cleavage at a cryptic site in the pre-mRNA. Furthermore they demonstrated that the location of the aberrant cleavage site was defined by a newly created base pair between the mutation in the invariant loop of U5 and the sequence just upstream of the new cleavage site. Mutation

of position 5 and/or 6 of the U5 invariant loop can activate cleavage at an aberrant site by base-pairing with nucleotides upstream of the new cleavage site at positions -2 and/or -3, respectively, (Chapter 1,Figure 5)(Newman and Norman, 1992). This U5·exon base-pairing interaction is reminiscent of the EBS·IBS (exon binding site/intron binding site) interaction observed in Group II autocatalytic introns, where this interaction is important for defining the 5' splice site (Chapter 1, Figure 2)(Newman and Norman, 1992).

Does the U5-exon interaction influence cleavage site choice in the nuclear pre-mRNA splicing pathway? Notably, mutations in the invariant loop do not activate the aberrant cleavage site if the 5' splice site is not mutagenized; rather, activation of aberrant sites is dependent on the presence of a mutagenized 5' splice site (Newman and Norman, 1992). Since the exon sequence adjacent to the 5' cleavage site is not highly conserved, it seems unlikely to contain sufficient information to constitute a conserved binding site in a fashion analogous to the U1·5' splice site or U2·branch point recognition sequence. However, the nucleotides in the invariant U5 loop whose mutations activate cryptic cleavage sites are normally uridines, prompting the suggestion that U5 can base pair promiscuously with the exon sequence (Newman and Norman, 1992).

Site-directed crosslinking of U5 snRNA to the 5' splice site. When a site-specific crosslinking agent was placed two nucleotides upstream from the cleavage site (position -2), U5 could be crosslinked to this pre-mRNA when incubated in a wild-type mammalian splicing extracts (Wyatt, et al., 1992). Remarkably, the

major crosslink mapped to the position in the U5 invariant loop (position 5, Figure 1) shown genetically, under the special conditions described above, to base pair with the same exon position in yeast. When the site-specific crosslinking agent was placed one nucleotide away at the exon nucleotide adjacent to the 5' cleavage site (position -1), the U5 invariant loop still crosslinked to the pre-mRNA. In this case, the crosslink mapped to an invariant loop nucleotide (position 4, Figure 1) adjacent to the one described above. Curiously, although the two site-specific crosslinks are in the same register, the kinetics of the two crosslinks are different (Sontheimer and Steitz, 1993). U5 first crosslinks to the penultimate exon position (position -2) and later to the ultimate exon position (position -1). These observations indicate that U5 is normally in close proximity to the exon sequence just upstream of the 5' cleavage site; however, the kinetics suggests that the U5 snRNA interacts in two different modes with the pre-mRNA.

Can U5 loop mutations increase cleavage at the aberrant site activated by actin intron position 5 mutations? As mentioned above, Newman and Norman (1991) demonstrated that mutations in the invariant loop of U5 could activate cleavage at aberrant sites if the cyh2 intron is mutated at position 1. The only other set of conditions, in yeast, that activate cleavage at aberrant sites is when either the act1 or rp51A intron is mutated at intron position 5 (Parker and Guthrie, 1985; Jacquier, et al., 1985). Thus, in the first case, an intron position 1 mutation combined with a U5 snRNA mutation can activate aberrant cleavage sites while in the

second case an intron position 5 mutation on its own can activate aberrant cleavage sites. Curiously, intron position 5 mutations in the cyh2 intron do not activate aberrant cleavage sites on their own or when combined with U5 snRNA point mutations (Newman and Norman, 1991).

Since the U5 invariant loop had been observed to crosslink to the 5' splice site in two different configurations prior to the first catalytic step, I reasoned that a different set of loop mutants might increase cleavage at the aberrant site activated by actin intron position 5 mutations. For this reason, I transformed the invariant loop library mutated at positions 2-8 of the invariant loop (Figure 1) (generously provided by A. Newman) into a cup1Δ strain carrying the ACT1-CUP1 reporter for aberrant cleavage activated by the intron G5C mutation. I then selected for transformants with increased copper resistance (Figure 2). I isolated many U5 alleles which increased cleavage at the aberrant cleavage site.

Figure 3A depicts the sequence of U5 snRNA activators of aberrant cleavage isolated in my screen. All of these suppressors are mutated in the invariant loop at either position 5 or 6 or both. At either position 5 or 6, the invariant U has been mutated to C (U5C or U6C). Notably, each of these mutations strengthens the possible base-pairing interaction between U5 and positions just upstream of the aberrant cleavage site [(-3)G(-2)G] (Figure 3B). Interestingly, the strongest U5 activators of the aberrant cleavage site were mutated at both positions 5 and 6 of the U5 invariant loop (U5C,U6C); in addition, the A at position 8 of the invariant loop was changed to a pyrimidine (Figure 3B). Thus, the mechanism of activation of

aberrant cleavage by the new U5 mutants is similar to that by the suppressors isolated by Newman and Norman (1991)(Chapter 1, Figure 5). This result was surprising since similar U5 loop mutations did not activate aberrant cleavage when the cyh2 intron was mutated at position 5 (Newman and Norman, 1991).

Why didn't the U5 mutants activate cleavage at an aberrant site in CYH2? The inability of the U5 mutants to activate cleavage in the presence of a wild-type CYH2 5' splice site (Newman and Norman, 1991) indicates that a normal interaction between the wild-type 5' splice site and some other, as of yet unknown, factor must be disrupted in order for U5 mutants to activate aberrant cleavage sites. Interestingly, the G5C mutation in the cyh2 intron, in contrast to the act1 intron, does not appear to have an effect on the splicing of this pre-mRNA. This mutation does not activate cleavage at an aberrant site and does not appear to decrease the efficiency of the first catalytic step (Newman and Norman, 1991). Thus, perhaps the G5C mutation is more disruptive in the act1 intron than in the cyh2 intron.

The only U5 mutant tested with the G5C cyh2 intron was the U5-U6C mutant (Newman and Norman, 1991). In the context of the cyh2 position 1 mutations, this change in the invariant U5 loop presumably activates aberrant cleavage because positions 5 and 6 of the loop (U5C6) base pair with the GA dinucleotide upstream of the newly activated aberrant cleavage site. This same mutation (U6C) was isolated in my screen to identify U5 loop mutations that increase cleavage at the aberrant site activated by the act1 G5C

mutation, presumably because position 6 of the loop (U6C) can base pair with the G at position -3 upstream of the aberrant cleavage site. Notably, this mutation is one of the weakest U5 loop mutants isolated in the screen. Therefore, perhaps "stronger" U5 loop mutants might be able to increase cleavage at aberrant sites in the cyh2 intron position 5 background.

These two sets of observations raise some interesting questions. Perhaps the ability of U5 loop mutants to activate aberrant cleavage sites is dependent upon (1) the phenotype conferred by the intron mutation on an unknown 5' splice site interaction, (2) the strength of the base-pairing interaction formed between the invariant loop of U5 and the exon sequence upstream of the new cleavage site, and (3) the nucleotide at position 8 of the loop. (Recall that in my screen the strongest U5 loop activators isolated also contained a pyrimidine at position 8.)

One interesting possibility is that the 5' splice site interaction which is being disrupted is actually the U1·5' splice site interaction. An in vitro purification assay has shown that pre-mRNAs mutated at intron position 1 bind less U1 snRNA than pre-mRNAs mutated at position 5 (Ruby and Abelson, 1988). Thus, perhaps the ability of the U5 mutants, isolated in my screen, to activate aberrant cleavage will be greater in the context of the actin intron position 1 mutations than in the presence of actin intron position 5 mutations.

How are cryptic cleavage sites chosen? In order to confirm their hypothesis that a U5 invariant loop·exon base-pairing

interaction was sufficient to define an aberrant cleavage site, Newman and Norman (1992) made a set of three point mutations in the invariant loop, at either position 5 or 6. In each case the mutagenized U5 snRNA activated cleavage at an aberrant cleavage site if the normal 5' splice site contained the G1A mutation. The site of the aberrant cleavage site, as predicted by their model, was always located just downstream of a new base-pairing interaction between the invariant loop and the pre-mRNA. However, in each case, cleavage was activated at only one aberrant site, even though there were several dinucleotides in the pre-mRNA sequence which were complementary to positions 5 and 6 of the U5 invariant loop. Thus, in addition to the sequence upstream of the new cleavage site there are other determinants involved in choosing the location of the new aberrant site.

In Table 1, I have listed the possible sites which could have been activated by each of the new U5 loop mutations. In each case, the site which is chosen is the site which most resembles a normal 5' splice site; that is, the sequence that could presumably form the most base pairs with U1 snRNA. These observations raise the possibility that U1 snRNA acts in conjunction with U5 snRNA to choose the aberrant cleavage site. Thus, conceivably U1 normally binds first to the pre-mRNA, but if this interaction is unstable, U1 and U5 may then act together to choose another 5' splice site by binding to the region with the greatest complementarity to both of the snRNAs. Consistent with this model is the observation that the first U5 site-specific cross-link to the pre-mRNA appears right after the U1 psoralen cross-link to the pre-mRNA and both peak and

diminish with very similar kinetics (Wyatt, Sontheimer and Steitz, 1992; Wassarman and Steitz, 1992).

Can U1 snRNA and U5 snRNA interact to define a 5' splice site? Consistent with the preceding hypothesis is the observation that, in my screen to isolate U5 mutations that increase cleavage at an aberrant site, the sequence downstream of the aberrant cleavage site (GUCUGG) resembles a 5' splice site (GUAUGU). Thus, in this situation, as in those described above, perhaps U1 and U5 cooperate to choose the location of the aberrant cleavage site. If this is true, I would make the following predications: (1) Mutations in the region downstream of the aberrant cleavage site that decrease complementarity to U1 should decrease activation of aberrant cleavage by U5 mutants. (2) Mutations in U1 snRNA which decrease complementarity to the normal 5' splice site and increase complementarity to the sequence downstream of the site of aberrant cleavage should increase activation of aberrant cleavage by U5 mutants. (3) Mutations in U1 snRNA which increase complementarity to the normal 5' splice site and decrease complementarity to the sequence downstream of the aberrant splice site should decrease activation of aberrant cleavage by U5 mutants. (4) Lastly, it should be possible to predict the site of activation of aberrant cleavage sites by different U5 loop mutations in the actin pre-mRNA.

Can U6 snRNA and U5 snRNA interact to influence the site of 5' cleavage? All the experiments discussed so far have implicated a possible interaction between U1 and U5 snRNAs in

choosing the 5' splice site. However, as described in Chapter 5, I have also genetically identified U6 snRNA mutations which affect the relative amount of cleavage at the normal and aberrant cleavage sites. Thus, both U5 and U6 snRNA mutants have been demonstrated to affect the amount of cleavage activated at the aberrant site by intron position 5 mutations. In addition, as alluded to earlier, while the first U5 site-specific crosslink follows similar kinetics to the U1 crosslink (Wyatt, et al., 1992), the second U5 crosslink only appears as the earlier crosslinks diminish (Sontheimer and Steitz, 1993). This second crosslink peaks at the time that the U6 crosslink and splicing intermediates appear. Thus, the crosslink results are consistent with the hypothesis that U1 and U5 first identify the 5' splice site region and then U5 interacts with U6 to determine the site of cleavage.

Interestingly, as proposed in Chapter 5, one determinant of the site of the actin aberrant cleavage could be the formation of fortuitous base-pairing interactions between U6 and the 5' splice site. However, such a base-pairing interaction cannot be used to explain the location of the aberrant cleavage site activated by position 5 mutations in the rp51A intron. On the other hand, the sequence just upstream of the aberrant cleavage sites in both introns could form either G·U or A·U base pairs with the wild-type uridines at position 5 and 6 of the U5 invariant loop. Thus, perhaps both U5 and U6 snRNAs are determinants in choosing the site of aberrant cleavage.

In order to compare the roles of U5 and U6 in determining the cleavage site, I transformed both a wild-type U6 strain and a U6-

A45G,U46A strain with the G5C or G5A ACT1-CUP1 reporter for aberrant cleavage along with two of the U5 loop mutants (U5-U5C,U6C,A8U and U5-U5C,A8U) isolated in my screen for U5 loop mutants that increase aberrant cleavage. The U6-A45G,U46A mutations increase complementarity to positions 7 and 8 of the actin intron. As described in Chapter 5, this double mutation decreases cleavage at the aberrant 5' splice site and increases cleavage at the normal 5' splice site. As demonstrated in Table 2, U5-U5C,U6C,A8U is a stronger activator of aberrant cleavage than U5-U5C,A8U, presumably because this mutant can form two G·C base pairs with the exon sequence upstream of the aberrant site, while the U5-U5C,A8U can form one G·C and one G·U base pair with the same sequence.

Interestingly, cleavage at the aberrant site, as inferred by assaying copper resistance, appears to reflect the effects of both the U5 and U6 mutants. For example, as summarized in Table 2, the presence of U5-U5C,U6C,A8U increases the copper resistance of a wild-type U6 strain carrying the G5C reporter for aberrant cleavage about three-fold, while a U6-A45G,U46A strain carrying the same reporter decreases copper resistance about three-fold. When the U6-A45G,U46A strain is transformed with the U5-U5C,U6C,A8U mutant, then copper resistance of the strain carrying the G5C reporter for aberrant cleavage is the same as the copper resistance of a wild-type strain transformed with the same reporter. Thus, it appears that the three-fold increase in aberrant cleavage conferred by the U5 mutation is negated by the three-fold decrease in aberrant cleavage conferred by the U6 mutation.

The preceding observations suggest that when intron position 5 is mutated, U5 and U6 each contribute equally to the decision of where to cleave. Since, the second U5 crosslink appears before the U6 crosslink, one interpretation of these data is that U5 chooses the site of cleavage and then directs U6 to the cleavage site. Although my new genetic results do not resolve whether the two snRNAs act consecutively or sequentially, they do indicate that neither snRNA plays a dominate role in choosing the site of cleavage.

How many times does U5 interact with the 5' splice site? Does U5 interact with both of U1 and U6 snRNAs? Notably, as discussed earlier, U5 has been crosslinked to two different nucleotides of the 5' splice site with different kinetics. Interestingly, the position of the early crosslink corresponds with the base-pairing interaction proposed by Newman and Norman (1991; 1992) and follows kinetics similar to the U1 crosslink, while the position of the second crosslink follows kinetics similar to the U6 crosslink (Wyatt, et al., 1992; Sontheimer and Steitz, 1993). Thus, the crosslink data support the notion that U5 might interact with the pre-mRNA twice, once with U1 snRNA and once with U6 snRNA.

Crosslinks between the U1, U5 and U6 snRNAs and the 5' splice site region of the pre-mRNA indicate that at some point in time the three snRNAs are in close proximity to the pre-mRNA. A comparison of the kinetics of the different crosslinks as well as the appearance of splicing intermediates and products suggests possible roles of these snRNAs in splicing. However, the crosslink studies do not address the functional significance of the physical interactions. In

contrast, genetic suppression experiments implicate U1, U5 and U6 in the function of 5' splice site selection. Mutations in each of these snRNAs have been identified that increase cleavage at the aberrant site activated by actin intron position 5 mutations, as reflected by the increase in copper resistance of strains harboring an ACT1-CUP1 G5C reporter for aberrant cleavage.

Primer extension analyses have proven very powerful in defining the roles of both U1 and U6 in 5' splice site selection. For example, compensating changes in yeast U1 snRNA, which increase cleavage at the aberrant site, decrease precursor-mRNA (pre-mRNA) levels and increase both normal and aberrant mature mRNA levels by similar amounts; consequently, the ratio of cleavage at the aberrant site to the normal site does not change. Thus, the observed increase in aberrant cleavage is a result of the overall increase in the efficiency of the first catalytic step resulting from the restoration of the U1·5' splice site base pairing interaction. In contrast, a mutation in U6 (U6-C43G) which increases cleavage at the aberrant 5' splice site does not affect, or only slightly increases, pre-mRNA levels; rather it increases aberrant mRNA levels and decreases normal mRNA levels so that the ratio of cleavage at the aberrant site to cleavage at the normal site changes (see Chapter 5). Thus, the U6 suppressor does not affect the efficiency of the first step, rather it alters the relative amount of cleavage at each of the splice sites.

Similar primer extension analyses can now be used to characterize the mechanism of suppression of the U5 snRNA mutations that increase aberrant cleavage. Thus, it remains to be

seen whether the U5 mutations decrease precursor levels, as do the U1 snRNA suppressors, or change the ratio of normal to aberrant mRNA, as do U6 snRNA suppressors, or both. These studies should indicate whether U5 plays a role in the selection of the region of the 5' splice site, like U1, or plays a role in choosing the exact cleavage site, like U6, or both.

Effects of snRNA mutations on competing 5' splice sites. In Chapter 2, I introduced the first of several 5' splice site competition assays. In this assay, the competing 5' splice sites are located only 13 nucleotides apart. All possible point mutations were introduced into all six nucleotides of the highly conserved intron portion of the 5' splice site. Each of these mutations decreased or abolished the ability of the mutated site to compete with the wild type site for cleavage. At the time these studies were being conducted, U1 snRNA was the only factor known to interact with the 5' splice site. Consequently, it was assumed that when the 5' splice site was mutated, thus reducing complementarity to U1, cleavage at the mutated site was prevented because U1 no longer bound. I was very surprised when compensating changes in U1 did not restore cleavage at the mutant site (see Appendix 2). In Chapter 3, I presented several hypothesis as to why the U1 compensatory changes did not restore cleavage at the mutated site. Interestingly, in one model ("splice site region versus cleavage site location"), I proposed that the inability of a mutated site to compete against the wild type site reflects a compromised interaction between the mutated site and a factor distinct from U1.

In light of the experiments described in this chapter, likely candidates for such a factor include U5 and U6 snRNAs. Interestingly, Jonathan Staley in our laboratory has recently demonstrated that mutations in both U5 and U6 snRNAs can activate cleavage at the mutated 5' splice sites. Preliminary experiments indicate that a 5' splice site mutated at intron position 6 can be suppressed by a compensating mutation at position 47 of U6. This observation confirms the base-pairing interaction proposed in Chapter 5, and indicates that this 5' splice site competition assay is monitoring determinants that choose the site of cleavage.

Interestingly, U5 loop mutations also increase cleavage at the mutated upstream 5' splice site. In contrast to the observations of Newman and Norman, mutations at position 4 of the invariant U5 loop are sufficient to activate cleavage at mutated upstream 5' splice sites; mutations at position 5 of the loop do not seem to activate cleavage at the mutated site. Preliminary experiments indicate that this suppression is the result of the formation of a new base-pair between the U5 loop position 4 and the last position of the exon (-1). Thus, in this case, it appears that increasing the complementarity between position 4, but not 5, of the invariant loop can affect 5' splice site selection. Strikingly, this new base pair is the same as the late U5 crosslink, which peaks at the time that splicing intermediates appear.

To summarize, the following observations raise the exciting possibility that the 5' splice site competition assay is monitoring the later decision of cleavage site choice rather than the early step of defining the region of the 5' splice site and commiting the pre-

mRNA to the splicing pathway: (1) restoration of complementarity between U1 and the 5' splice site does not activate cleavage at a mutated site; (2) restoration of complementarity between U6 and the 5' splice site does activate cleavage at a mutated site, and (3) cleavage of a mutated 5' splice site is increased by creating complementarity between U5 and the 5' exon by a pairing that is analogous to the major U5 late crosslink (Sontheimer and Steitz, 1993).

Summary. Three of the 5 snRNAs, U1, U5 and U6, have been implicated in interactions with the 5' splice site. U1 is involved in the initial recognition of the pre-mRNA by base-pairing with the 5' splice site. U6 later base pairs with a subset of the nucleotides of the 5' splice site and acts to determine the exact location of cleavage. The challenge remains to determine the role of U5 in splice site recognition. Interestingly, the base-pairing interactions between U1 and the 5' splice site and U6 and the 5' splice site are mutually exclusive. Thus, U1 must be displaced from the 5' splice site before U6 can form Watson-Crick base pairs with the same sequence. Thus, as shown in Figure 4, one function of U5 snRNA or its associated proteins may be to maintain definition of the 5' splice site while U1 leaves and U6 arrives.

Figure 1. U5 snRNA is in close proximity to the 5' splice site. The nucleotides of the U5 loop are numbered 1-9 and are shown in close proximity to the exon portion of the 5' splice site consensus sequence.

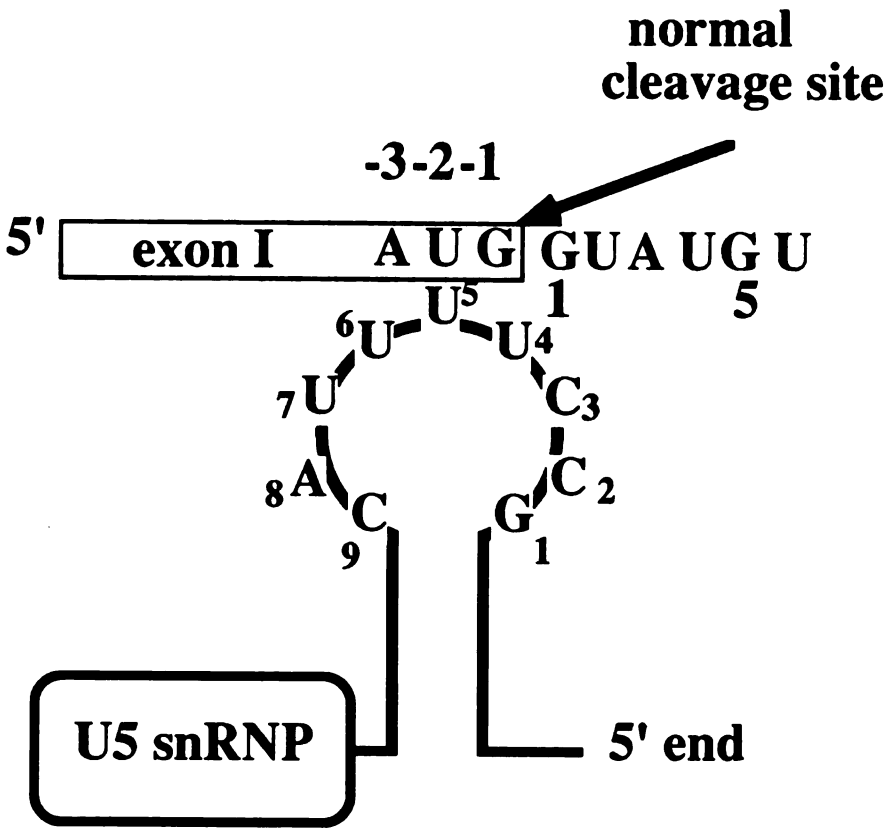
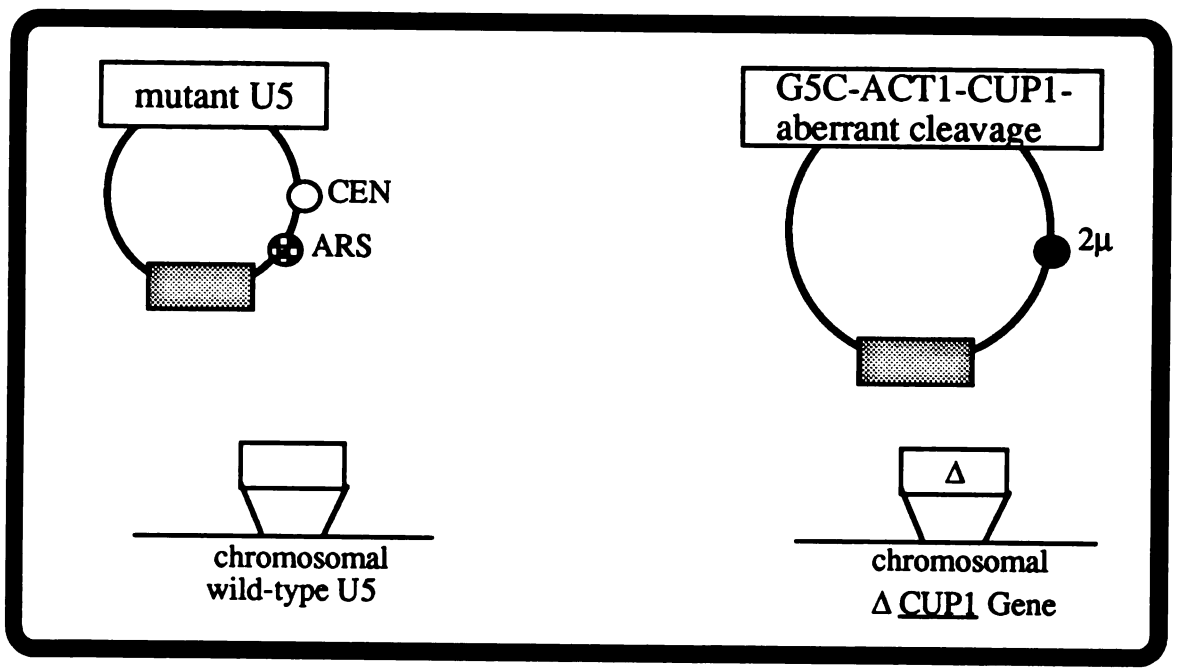

normal
cleavage site
-3-2-1
5'
exon I
A U G
G U A U G U
1
5
U5
6U
U4
7U
C3
8A
C
C 2
G
9
1
U5 snRNP
5' end

Figure 2. Genetic scheme used to isolate U5 snRNA point mutations that increase cleavage at the aberrant cleavage site. A bank of U5 snRNA point mutations was transformed into a haploid strain that was deleted for the chromosomal copy of CUP1. The introduction of the library made the strain heterozygous for U5 snRNA. The strain carried the G5A ACT1-CUP1 reporter for aberrant cleavage. U5 snRNA mutations were screened for based on their ability to increase copper resistance as a result of cleavage at the aberrant 5' splice site.

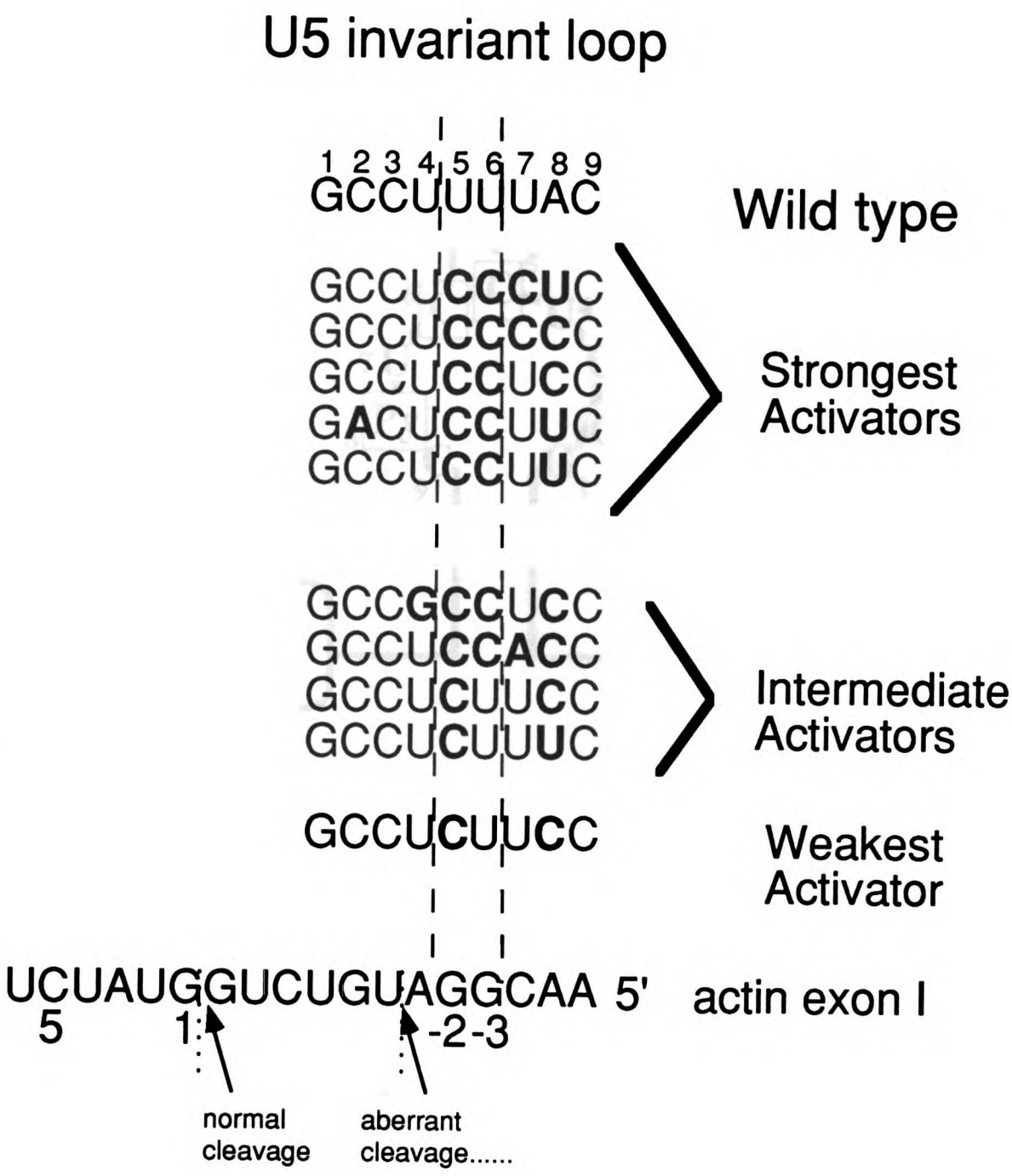
transform in U5 loop
mutant library

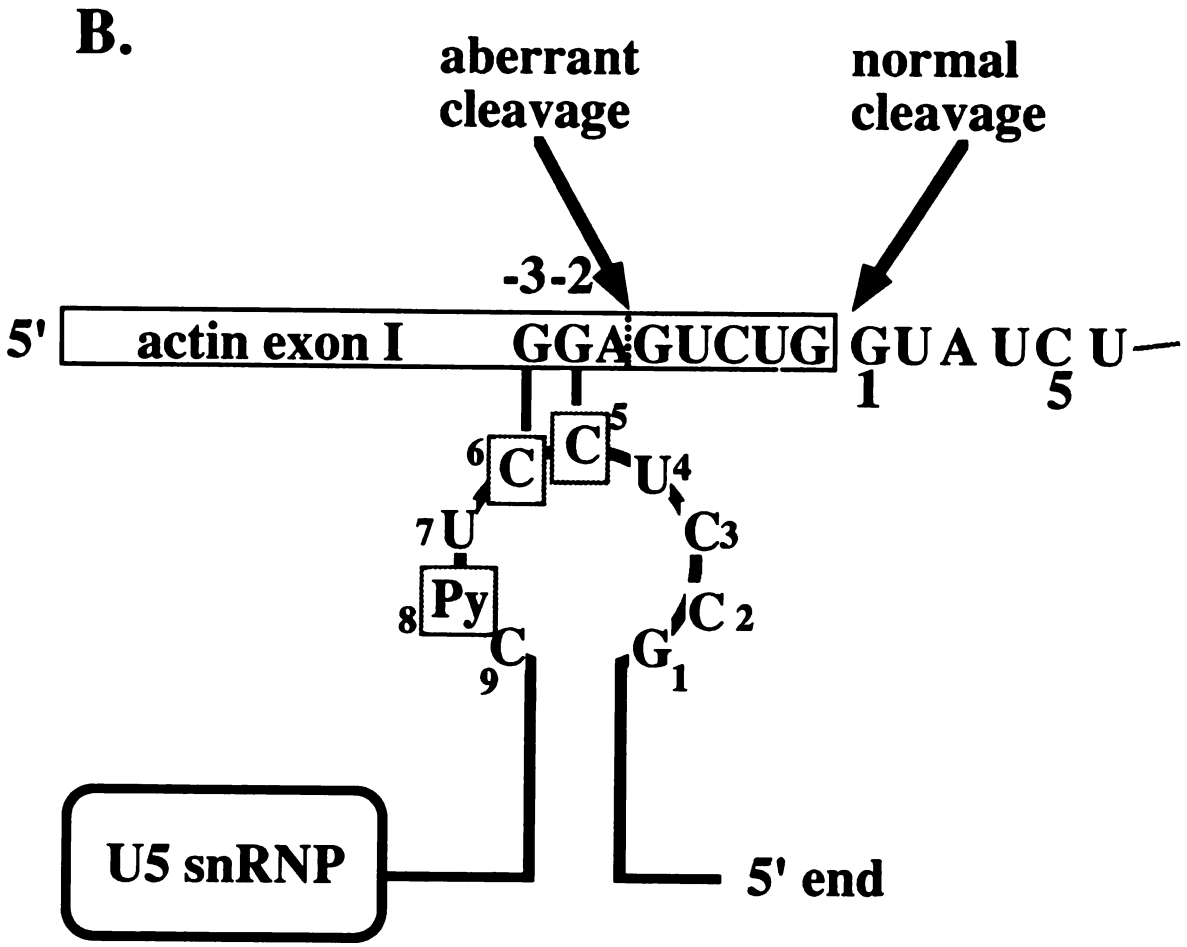


Select for increased
copper-resistance

Figure 3. Summary of U5 snRNA point mutations that increase cleavage at the aberrant cleavage site. **(A)** Listed are the U5 snRNA loop mutations isolated in the genetic screen depicted in Figure 2. The mutants are grouped according to their ability to increase cleavage at the aberrant site as reflected by copper resistance levels. The sequence of the actin G5C reporter is shown at the bottom of the figure. **(B)** A model for increased activation of aberrant cleavage by U5 loop mutants. The strongest activators of aberrant cleavage are mutated at positions 5,6 and 8 of the invariant loop.

A.


U5 invariant loop
1 2 3 4 5 6 7 8 9
GCCUUUUAC
Wild type
GCCUCCCUC
GCCUCCCCC
GCCUCCUCC
GACUCCUUC
GCCUCCUUC
Strongest Activators
GCCGCCUCC
GCCUCCACC
GCCUCUUCC
GCCUCUUUC
Intermediate Activators
GCCUCUUCC
Weakest Activator
UCUAUGGUCUGUAGGCAA 5'
actin exon I
5
1
-2-3
normal cleavage
aberrant cleavage.....


B.
aberrant cleavage
normal cleavage
-3-2
5'
actin exon I
GGAGUCUG
GUA UC U—
1
5
6C
C5
U4
7U
C3
8Py
C2
C
G1
9
U5 snRNP
5' end

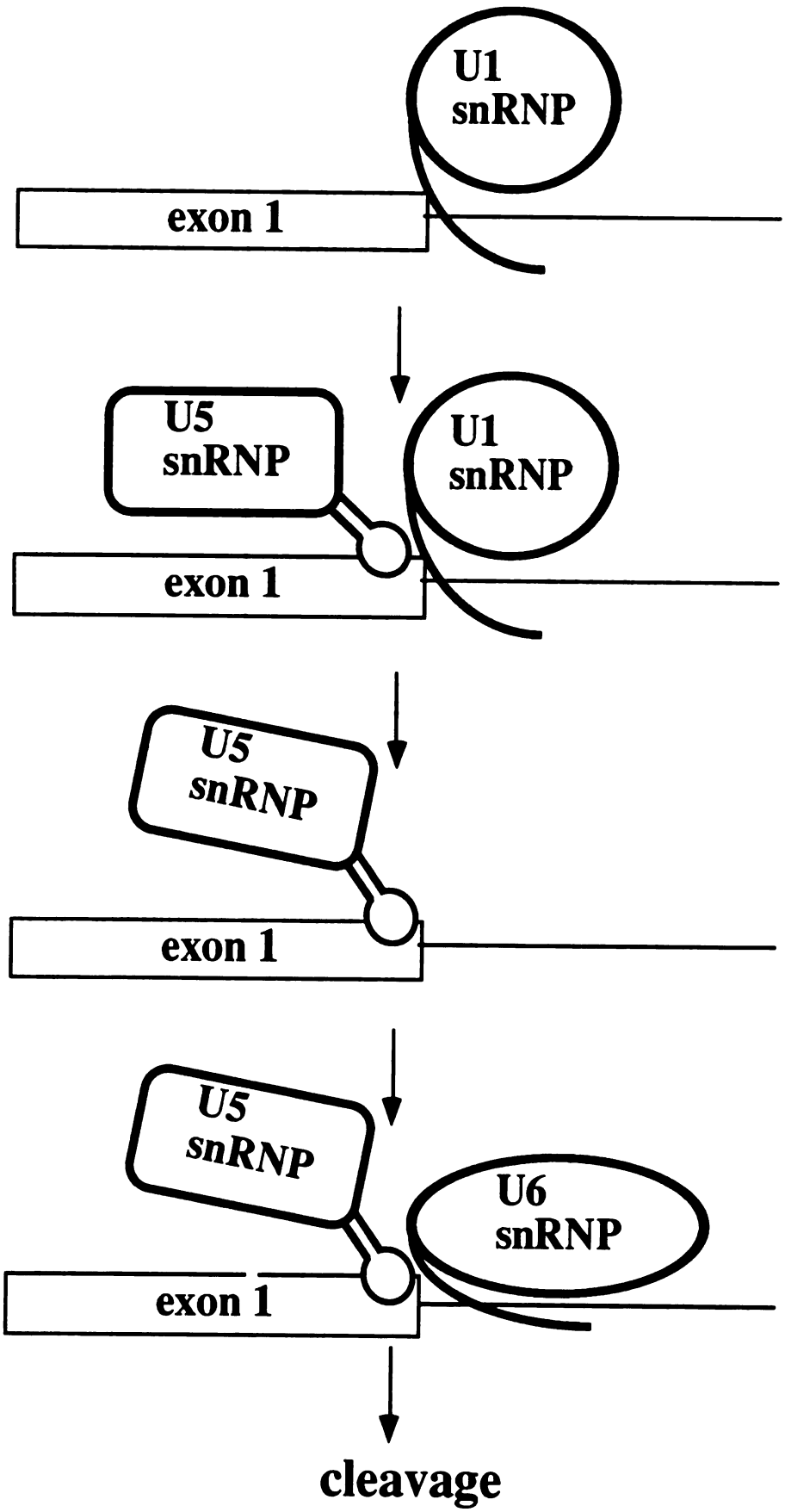

U1
snRNP
exon 1
U5
snRNP
U1
snRNP
exon 1
U5
snRNP
exon 1
U5
snRNP
U6
snRNP
exon 1
cleavage

Table 1 Analysis of sites activated by U5 snRNA loop mutations when the cyh2 intron is mutated at intron position 1.

mutant U5 snRNA*	aberrant site activated†	alternative site(s) not activated
CUU**A**	**U**A**U**/**GUA**GU**U**	**UAG**/U**U**CCA**U**
CU**C**U	**AGG**/UC**A**C**GU**	**AG**U/UCCAUA **AG**A/UAUGUA **AG**A/**G**GUCAG
CUU**C**	**GAG**/**GU**CACG	**GA**U/A**U**GUAG

* Bold nucleotides are mutated positions in the U5 snRNA invariant loop.

† Bold underlined letters = nucleotides of the pre-mRNA that are complemenatary to either U1 or U5.

Table 2 Effects of U5 and U6 snRNA mutations on copper resistance generated by cleavage at the aberrant cleavage site activated by the actin intron G5C or G5A mutations.

Copper resistance of G5C aberrant cleavage reporter			
	U5 snRNA		
U6 snRNA	GCCUUUUAC	GCCUCCUUC	GCCUCUUUC
U6-WT	.05+++.1~	.18+++	.18~
U6-45G46A	.025~	.05+++.1~	.05+
U6-C43G	.1+++	.18+++	.18~
U6-C43U	.05++	.18+++	.18~

Copper resistance of G5A aberrant cleavage reporter			
	U5 snRNA		
U6 snRNA	GCCUUUUAC	GCCUCCUUC	GCCUCUUUC
U6-WT	.5+	.1+++.18~	.18+++.25~
U6-45G46A	.5+++	.5+++	.5+++
U6-C43G	.1~	.1~	.1~
U6-C43U	.25+++.5~	.18+	.18+++.25~

Epilogue

How is the 5' splice site region recognized?

How is the 5' exon-intron boundary chosen?

Amazingly, introns are recognized and correct cleavage sites are chosen based solely on information provided by three short consensus sequences. Consequently, rather than starting with a particular splicing factor and trying to determine its role, I chose to "view" splicing from the intron's perspective. That is, my approach has been to analyze the phenotypes conferred by mutations in the intron consensus sequence and then to use this information to design genetic screens to identify trans-acting components.

I chose to focus my analyses on the 5' splice site sequence since, as described in Chapter 1, this sequence is involved in several steps of the splicing process: (1) identification of the 5' splice site, (2) commitment of the pre-mRNA to the splicing pathway, (3) definition of the exact cleavage site, and (4) determination of whether the splicing apparatus should progress through the second step of splicing after inspection of the branched lariat intermediate.

Determinants of 5' splice site selection. As described in Chapter 2, in order to identify the nucleotides of the 5' splice site that are presumably important for splice site recognition, I set up a 5' splice site competition assay. In this assay the competiting 5' splice sites are separated by 13 nucleotides. Mutations were introduced into all of the conserved nucleotides of the upstream 5' splice site. Three interesting observations were made: (1) Mutations in the most highly conserved nucleotides of the intron portion of the consensus sequence, which in several cases have been demonstrated to base pair with U1 snRNA, almost totally abolished recognition of the mutated 5' splice site. (2) Unlike mammals,

complementarity of U1 snRNA with the nucleotides in the exon portion of the 5' splice site is not an important determinant of 5' splice site choice. (3) All changes at intron position 4 allowed competition with the wild-type site (C>G>A). This result is interesting since the 5' end of the yeast U1 and the mammalian U1 are identical, yet the intron consensus sequences differ at position 4; in yeast this position is almost invariably unpaired. The competition results indicated that this is the favored configuration.

The most likely explanation for the observation that mutations in the intron portion of the splice site sequence caused step 1 defects that decreased or abolished spicing at the competing mutated splice site was that the mutant 5' splice sites couldn't compete with the nearby wild type site for U1. A straightforward way to test this hypothesize was to determine whether compensating changes in U1, predicted to restore base pairing with a mutated site, could increase use of the competing mutated site. However, as discussed in Chapter 3, the restoration of complementarity between U1 and the mutated 5' splice site by compensating mutations in U1 did not increase cleavage at the mutated site. Thus, the inability of the mutated upstream 5' splice sites to compete with the nearby wild-type 5' splice site is not (simply) a result of their decreased complementarity to U1. In addition, as discussed in Chapter 3, these observations raise the question as to what step of splicing this competition assay is monitoring. And, are other factors involved in 5' splice site selection before or after U1 snRNA binds?

Chapter 4 reviews my attempts to further characterize the interaction of U1 with the 5' splice site. By constructing several 5' splice site competition assays, I was able to conclude that the process of 5' splice site selection in yeast, like that in mammals, is not solely determined by a base-pairing interaction between U1 and the 5' splice site. The sequence of the 5' splice site, the sequence flanking the highly conserved 5' splice site and the positioning of the 5' splice site relative to the branch point recognition sequence all appear to influence 5' splice site selection. Recent observations about the role of proteins in mammalian 5' splice site selection (see Chapter 1) and my observations in yeast (Chapters 3 and 4), make several predictions about the role of proteins in 5' splice site selection and can be used to design genetic schemes to identify such proteins in yeast.

The clearest conclusion that can be drawn form Chapters 2-4 is that the process of 5' splice site selection is complicated and involves many determinants including (1) splice site sequence, (2) flanking sequence, (3) proximity to branch point sequence and (4) complementarity to U1 snRNA. One way to rationalize this complexity is to consider that recognition of the 5' splice site sequence commits the pre-mRNA to the splicing pathway. Thus, the splicing apparatus must ensure that the correct substrate is chosen before proceeding through the ATP-dependent splicing pathway.

Choosing the site of cleavage at the 5' exon-intron boundary. Interestingly, as discussed in Chapter 1, the processes of recognizing the 5' splice site region and choosing the exact site

of 5' cleavage appear to be separable events. In order to identify specific factors involved in determining the location of cleavage of the 5' splice site, I conducted many unbiased genetic screens which were unsuccessful. However, surprisingly, the factors I was trying to identify turned out to be well characterized splicing factors (indeed, the first identified members of the spliceosome), the snRNAs.

A series of both yeast and mammalian in vitro crosslinking experiments (reviewed in Chapter 5) suggested that U6 snRNA is in the proximity of the 5' splice site during the splicing reaction (Sawa and Abelson, 1992; Wassarman and Steitz, 1992). In yeast splicing extracts, UV crosslinks were identified between U6 and either the actin pre-mRNA or the actin lariat intermediate. The crosslink to the lariat intermediate was between a region of U6 just upstream of the invariant ACAGAG sequence and intron positions 4 through 6 (Sawa and Abelson, 1992). In mammalian splicing extracts, a psoralen crosslink was mapped between the invariant ACAGAG sequence of U6 and a region of the pre-mRNA which contains the 5' splice site (Wassarman and Steitz, 1992). Interestingly, in the mammalian system, the appearance of the U6-pre-mRNA crosslink was concomitant with the appearance of splicing intermediates, implicating a role for U6 in 5' cleavage. Based on these different crosslinking experiments, two distinct U6·5' splice site base-pairing interactions were proposed.

As discussed in Chapter 5, these observations were exciting because they provided a framework for explaining the activation of aberrant cleavage sites by intron position 5 mutations. If a U6·5'

splice site interaction was important in defining the site of cleavage, then mutations at intron position 5 could result in the activation of aberrant sites because they decreased the stability of the normal U6·5' splice site base-pairing interaction. In order to test this hypothesis, I extended the region of complementarity between U6 and the actin 5' splice site so that the normal U6·5' splice site interaction would be strengthened. Under these conditions, activation of the aberrant cleavage site by intron position 5 mutations is diminished. Thus, these experiments confirmed that a base-pairing interaction consistent with that proposed based on the mammalian psoralen crosslink is likely to take place and furthermore, this interaction is an important determinant for the site of cleavage. In addition, these results might explain the conservation of a U at position 4 of the yeast 5' splice site consensus sequence. Thus, while intron position 4 can not base pair with the corresponding nucleotide in U1 snRNA, this position might be conserved because it is complementary to the corresponding nucleotide in U6 snRNA.

In addition to base-pairing with the 5' splice site, U6 snRNA has recently been demonstrated to base pair with U2 snRNA (Madhani and Guthrie, 1992) (reviewed in Chapter 4). This U6·U2 helix is located immediately downstream of the highly conserved and essential nucleotides of the ACAGAG sequence in U6 and immediately upstream of the branch point recognition sequence of U2. Thus by base-pairing with both U2 and the 5' splice site, U6 can serve the critical function of aligning the 5' splice site of the substrate with the branch point nucleotide, which is the attacking

nucleophile in the first chemical step (Query, Moore and Sharp, submitted).

As discussed in Chapter 6, under special conditions, mutations in U5 snRNA can also influence cleavage site choice. In particular, I have identified mutations in the invariant loop of U5 snRNA that increase cleavage at the aberrant site activated by intron position 5 mutations. These U5 mutants appear to increase aberrant cleavage by creating complementarity between the invariant loop of U5 snRNA and the sequence upstream of the aberrant cleavage site. Curiously, as discussed in Chapter 6, it is not yet clear whether U5 acts early along with U1 snRNA to define the 5' splice site region of the pre-mRNA or later with U6 snRNA to choose the exact location of cleavage or both. Interestingly, the base-pairing interactions between U1 and the 5' splice site and U6 and the 5' splice site are mutually exclusive. Thus, U1 must be displaced from the 5' splice site before U6 can form Watson-Crick base pairs with the same sequence. One function of U5 snRNA or its associated proteins may be to maintain definition of the 5' splice site while U1 leaves and U6 arrives.

In summary. The 5' splice site appears to be involved in several steps of the pre-mRNA splicing pathway. In particular, U1 and presumably other factors are involved in the initial recognition of the pre-mRNA by base-pairing with the 5' splice site. This interaction commits the pre-mRNA to the splicing pathway. Once the pre-mRNA has been committed to the splicing pathway, the spliceosome is faced with the challenge of correctly juxtaposing the

catalytic residues of the spliceosome so that the correct cleavage sites are selected. Notably, my work has played an important role in demonstrating that in pre-mRNA splicing, like in self-splicing introns, it is RNA-RNA interactions which serve to correctly juxtapose the 5' splice site and ensure that the correct site of cleavage is chosen at the 5' exon-intron boundary. Preliminary evidence suggests that one role of U5 snRNA or its associated proteins might be to link together these two steps of 5' splice site selection.

Appendix I

Assaying copper resistance

Chapter 2 describes the development of a new genetic system which can be used to study basic cellular mechanisms by fusing CUP1 to genes of interest. The CUP1 system is very sensitive and has a large range of detection. However, presently, this system is limited in that it is difficult to objectively quanitate copper resistance. The following section reviews assays presently being used to monitor copper resistance in addition to suggesting possible new quantitative assays.

I. Copper resistance assays

In order to maximize the range over which it is possible to monitor copper resistance, the ACT1-CUP1 fusion gene was expressed on a high copy (2μ) plasmid. Each cell transformed with this construct will contain at least 10 copies of the 2μ plasmid if it is a *cir+* strain (Rine, 1991; Broach and Volkert, 1991). In a population of cells, the number of 2μ plasmids in any individual cell reflects a Poisson distribution (Broach and Volkert, 1991). Thus individual cells in this group will have varying levels of apparent copper resistance due to differences in plasmid copy number. Consequently, a culture of cells will not possess a distinct growth threshold; rather, only a small proportion of the cells will grow when exposed to high copper concentration. This observation was taken into account when designing the following copper resistance assays

A. Qualitative methods

i. **Patch assay**

This is the easiest assay. Patch out 3-4 independent transformants on a plate deficient for the auxotrophic marker of the plasmid carrying the CUP1 reporter (as well as the auxotrophic markers of any other relevant plasmids). Allow the patches to grow for 1-2 days, and then replica the master plate to minimal complete plates with increasing concentrations of copper. Incubate the copper-containing plates for 3-7 days, and then score copper resistance. There is no reason to select for plasmids required to maintain copper resistance on copper plates. However, if other plasmids might lower the copper resistance of the strain by inhibiting expression of the reporter, then selecting for auxotrophic markers ensures that the plasmids are not lost.

The density of the patch will decrease before no growth can be detected as the strain is transferred to plates that contain increasing amounts of copper. Thus, it is a subjective decision to define the maximum copper concentration at which the strain can grow. However, it is possible to roughly compare the copper resistance of two different strains by observing the density of the patches at a particular copper concentration. For example, if strain A is resistant up to .25 mM copper and strain B is resistant up to 1.5 mM copper, on .18 mM copper plates, the density of strain B will be greater than that of the strain A.

Note that, in addition, the actual copper concentration of the plates will vary slightly from batch to batch. Thus, in all of the copper resistance assays, it is important to compare the copper resistance of experimental strains to well-characterized strains.

ii. Frogging assay

This is also a relatively easy assay. Importantly, frogging can be more uniform than replica-plating, since a similar amount of cells in a similar volume is transferred to each copper-containing plate. First, 0.1 ml of sterile minimal media is added to each well of a sterile microtiter dish. Next, cells from either a colony or patch of the strain of interest are stirred into the microtiter wells using a sterile toothpick. A 48-prong metal frog is then used to replica the strains in the wells to minimal media plates that contain increasing amounts of copper. The cells should be transferred three times to three different plates that contain the same amount of copper in order to ensure that the correct amount of cells were transferred by the frogging device. Once the strains are plated, copper plates are incubated and scored as described above.

B. Potential quantitative assays

As mentioned above, defining the lowest copper concentration at which a particular strain can no longer grow can be a major problem. The following two assays describe potential methods for quantitating copper growth.

i. Percentage growth assay

The idea behind this assay is to define copper resistance as the highest concentration at which a pre-determined percentage of the transformants carrying the plasmids can grow. The percentage of cells growing in a replica-plated or frogged patch cannot be reliably determined. Thus, an alternative method is to grow up an overnight

culture of the strain of interest in minimal media to select for plasmids and then plate a dilution of cells to give 300-400 independent colonies onto a series of plates that contain increasing amounts of copper, as well as a plate that contains no copper. By tallying the number of cells on each of the copper plates and normalizing this value to the number of cells that can grow in the absence of copper, it is possible to determine the percentage of cells that can grow at a particular copper concentration.

While designing the percentage growth assay, I noticed that the size of the colonies growing on the copper plates decreased as the copper concentration increased. Thus, in the patch assay described above, the thickness of the patches decreases before disappearing because at high copper concentrations (1) the size of the colonies in the patches decreases and (2) only a small percentage of the cells will express the ACT1-CUP1 2μ plasmid at high enough levels to confer the appropriate copper resistance.

ii. Liquid growth assay

This assay is not very well developed yet, but has considerable potential. Sensitivity to copper can be detected in copper-containing liquid media in addition to solid minimal media plates; however, strains are resistant to lower levels of copper in liquid minimal media. Interestingly, at low copper concentrations, where strains which exhibit both high or low levels of copper resistance can grow, the OD600 of a culture of a strain which exhibits high copper resistance will be much greater than a strain that exhibits low copper resistance. This observation is consistent with the

observation (see patch assay) that it is possible to roughly compare the copper resistance of two different strains by examining the relative density of the patches at a chosen copper concentration. Presumably, the OD600 of the cells or the density of the patch reflects both the percentage of the cells which express the ACT1-CUP1 reporter at a high enough copy number to confer the necessary copper resistance as well as the size of the colonies at the chosen copper concentration. The advantage of measuring the OD600 of the cultures is that, in this way, it is possible to make an objective quantitative comparison, while in the patch assay it is a much more subjective qualitative comparison.

The first step in this assay is to grow up overnight cultures of the strains of interest and measure the OD600 of each culture. Next dilute each culture to a set OD (~0.1) in liquid minimal media in the presence and absence of a pre-determined copper concentration. Allow the cultures to grow for several days and then take the OD600 of each culture. Normalize the growth in the copper-containing media relative to the growth in the copper-lacking media. In this way it may be possible to quantitatively compare the copper resistance of a set of strains.

II. Copper resistance vs. copy number and promoter. As described in chapter 2, one of the goals considered when designing the ACT1-CUP1 gene fusion system was to design a system which could differentiate production of the gene fusion reporter product over a large range. Thus, the ACT1-CUP1 fusion gene was placed under the control of several different promoters either the

endogenous actin promoter or the strong constitutive ADH (alcohol dehydrogenase) or GPD (glucose-6-phosphate dehydrogenase) promoters, and maintained at different copy numbers in the cell by either integrating the fusion gene into the chromosome or placing it on a low (cen) or high (2μ) copy plasmid. The following table summarizes the copper resistance of ACT1-CUP1 wild-type gene fusions expressed in a wild-type strain.

expression source	promoter	copper resistance
integrated	GPD	.025 mM
cen	actin	0.10 mM
cen	GPD	0.25 mM
2-micron	ADH	0.50 mM
2-micron	GPD	1-2 mM

The copper resistance of a Δcupl strain is .013mM copper. Thus, when integrated into the chromosome, even under the control of the strong GPD promoter, the range of detection of the ACT1-CUP1 is only about two-fold. In contrast, when expressed on a high copy plasmid under control of the strong GPD promoter, the range of detection is about 100-fold. Thus, in the latter situation, the large range of detection coupled with the sensitivity of CUP1 gene fusions, results in an increase in the ability to differentiate between phenotypes conferred by mutations in a particular CUP1 fusion gene system.

III. Caveats about the Δcup1 strain

As described in Chapter 2, the original CUP1 deletion was marked by URA3. No information is available on how the original CUP1 deletion was created. In order to make the strain ura3$^-$, I made a small deletion in the coding sequence of the URA3 gene marking the CUP1 deletion. Consequently, this strain has two non-functional ura3 genes: one at the normal chromosomal copy of URA3 (ura3-52) and another at the CUP1 locus marking the CUP1 deletion. A low background of URA3$^+$, cup1$^-$ cells arises, presumably from gene conversion between the two ura3 loci. Thus, sometimes when this strain is transformed with a plasmid marked by URA3, Ura$^+$ transformants are generated which do not contain the plasmid. To distinguish between reversion of one of the chromosomal ura3 copies and URA3$^+$ transformants, streak the cells on 5-FOA. Only URA3 plasmid-bearing strains, and not URA3 chromosomal revertants, should be able to grow on 5-FOA.

Appendix 2

Effects of prp and slu mutations on 5' splice site competitions

Splice site competition assays have proven to be useful tools for identifying proteins that interact with these sequences. As discussed in Chapter 1, the amount of SR proteins added to a splicing extract can influence cleavage at competing 5' splice sites in vitro (Krainer, et al., 1990; Zahler, et al., 1993). In addition, by examining the phenotypes conferred by mutant strains on splice site competition assays, it has been possible to genetically identify factors which interact with the splice sites. For example, 3' splice site competition assays have been very useful in identifying mutant alleles of splicing proteins that are defective in specific aspects of 3' splice site choice (Frank and Guthrie, 1992; Umen and Guthrie, unpublished).

In an attempt to identify factors which interact with the 5' splice site, 5' splice site competition reporters were introduced into a subset of strains known or suspected to have splicing defects (**p**re-m**R**NA **p**rocessing and **s**ynthetic **l**ethal with **U5** strains). Each strain was transformed with three of the 5' splice site competition reporters (introduced in Chapter 2): (1) the wild-type/wild-type competition (2) the U4A/wild-type competition and (3) the G5C/wild-type competition.

The wild-type/wild-type competition should test if 5' splice site choice is affected by the location or context of the 5' splice site (or mRNA stability- see below.) The G5C/wild-type and U4C/wild-type competition were chosen because each of these mutant 5' splice sites can compete with the wild type site to some extent, thus it is possible to monitor either increases or decreases in cleavage at the mutant site. Moreover, while intron position 5 is

normally complementary to U1, position 4 is not normally complementary to U1, and the U4A mutation actually increases complementarity between U1 and the 5' splice site. As discussed in Chapter 2 and 3, position 4 probably plays a role in splicing independent from recognition by U1. I hoped that by comparing the phenotypes conferred by splicing mutants on the different reporter competitions it would be possible to identify factors involved in 5' splice selection.

Each of the competition reporters was transformed into the strains listed in the tables below. It was not possible to analyze splicing by copper resistance assays since these strains did not carry deletions at the chromosomal CUP1 locus. Rather, total RNA was isolated from each of these strains, grown at their permissive temperature (usually 25°C), and primer extended. Phosphorimaging techniques were used to quantitate the amount of mRNA generated as a result of cleavage at each of the competing 5' splice sites. Most of these mutant strains do not confer a detectable splicing defect at a permissive temperature; however, often when two splicing mutants are combined a splicing defect will be observed at the otherwise permissive temperature (Frank and Guthrie, 1992). Thus, I first decided to examine the splicing phenotypes conferred by the mutant strains at their permissive temperatures.

The data are summarized in Table 1. Each of the values given represents the amount of cleavage at the upstream (wild-type, U4A or G5C) site relative to cleavage at the downstream wild-type site. For example, in the wild-type strain cleavage at the wild-type upstream site occurs 1.9 times as often as cleavage at the

downstream site. However, when the upstream site is mutated to U4A or G5C, cleavage at this site is decreased five (.37/1.9)- or ten (.19/1.9)-fold, respectively. In Table 2, the numerical values have been replaced by arrows which indicate the amount of cleavage at the upstream site in the trans-acting mutant strain, relative to the amount of cleavage at the upstream site in the wild-type strain. For example, the prp5 strain decreases (↓) cleavage slightly at both the wild-type and the U4A upstream sites relative to the wild-type strain (1.4 versus 1.9 and .37 versus .26), but increases cleavage (↑) at the G5C upstream site relative to the wild-type strain about two-fold (.36 versus .19).

As can be seen in the first column of table 2, many of the mutant strains appear to affect the ability of the upstream wild-type site to compete with the downstream wild-type site. Does this phenotype reflect splicing at each of these sites? It is important to note, that in each of these competition assays, cleavage at the upstream site results in the generation of a stable in-frame pre-mRNA while cleavage at the downstream site results in the generation of an unstable out-of-frame pre-mRNA. Thus, if for any reason the strains being compared alter the stability of in-frame mRNA relative to out-of-frame mRNA, then this would result in a change in relative stability of the mRNAs generated by cleavage at each of the splice sites and might consequently be confused with a splicing phenotype. Interestingly, relative cleavage at each of the competing wild type sites has been observed to vary in different strain backgrounds (data not shown).

Since it was possible that changes in the ratio of cleavage at the upstream site relative to the downstream, in the mutant strains, was due to a change in the cell unrelated to splicing, the amount of cleavage at each of the upstream sites was normalized to cleavage at the wild-type upstream site. These values are summarized in Table 3. For example, in the prp5 strain cleavage at the wild-type upstream site is slightly decreased relative to the wild-type strain (1.4 versus 1.6). When the amount of cleavage at the U4C or G5C sites is adjusted for the decrease in apparent cleavage at this site, the decrease in cleavage observed at the U4A site is no longer significant while the increase in cleavage at the G5C site appears even more significant. In Table 4, like Table 2, the numerical values have been replaced by arrows which indicate the amount of cleavage at the normalized upstream site, in the mutant strain, relative to the amount of cleavage at the normalized upstream site, in the wild-type strain.

Although the data are quite complex, the splicing mutants can be grouped into three classes.

Class I: The prp3, prp7 and U5-U98C strains all decrease cleavage at the upstream site regardless of whether the site is mutated or not (Table 2). However, even when normalized for the decrease in cleavage at the wild-type upstream site (see above), the mutant sites are poorly recognized (Tables 3 and 4). Notably, the prp3 strain conferred a very strong phenotype. Preliminary experiments to test whether the splicing phenotype conferred by the prp3 strain could be suppressed by introducing a wild-type copy of PRP3 into the

cell gave conflicting results. Only some of the experiments indicated that the introduction of the wild-type allele suppressed the splicing phenotype. Thus, these experiments should be reproduced in order to resolve this discrepency.

Class II: The prp5, prp9 and prp11 strains all decrease cleavage at the wild-type upstream site and the U4A site, but increase cleavage at the G5C site (Table 2). Even when normalized for the decrease in cleavage at the wild-type site, in all three of the prp strains, cleavage is still increased at the G5C site. And, in the case of prp 11, cleavage is still decreased at the U4A site while in the prp5 and prp9 cleavage at the U4A site is similar to wild-type.

Why would cleavage be more efficient at a mutant 5' splice site (G5A) with decreased complementarity to U1, but less efficient with one (U4A) increased in complementarity to U1? Perhaps the U1/5' splice site interaction needs to be disrupted before U2 can bind. Interestingly, in vitro assays have demonstrated that U1 can bind and commit the pre-mRNA to the splicing pathway in the absence of ATP, but U2 is not assembled until ATP is added (Seraphin and Rosbash, 1989a). Curiously, U2 can bind to the pre-mRNA in the absence of ATP in extracts from U1-C4U mutant strains (Liao, et al., 1992). This mutation decreases complementarity between U1 and the 5' splice site. In addition, recent experiments suggest that the U1·5' splice site interaction must be disrupted in order for U2 to bind to the spliceosome (Konforti et al., 1993). Thus, one exciting interpretation of these experiment is that the ATP requirement for U2 addition is to mediate destabilization of the

U1/5' splice site interaction. These results are very exciting since recent biochemical and genetic experiments have shown that the PRP5, PRP9, PRP 11 PRP21 proteins interact (Legrain and Chapon, 1993; Legrain, et al., 1993; Ruby, et al., 1993). These strains all confer step 1 splicing defects and are required for U2 addition to the spliceosome after U1 has bound (Ruby, et al., 1993). Interestingly, PRP5, one of the proteins required for U2 addition is a member of the DEAD-box family and is a putative helicase (Dalbadie-McFarland and Abelson, 1990). Thus, the role of PRP5, and perhaps the PRP9, PRP11, PRP21 complex might be to disrupt the U1/5' splice site interaction to allow U2 to bind and splicing to proceed. According to this model, when any of these proteins are mutated, cleavage at the G5C site is increased because the U1/mutant 5' splice site interaction is easier to disrupt, allowing U2 to bind. In contrast, the U4A mutation, which increases complementarity between U1 and the 5' splice site, should make it more difficult to disrupt the U1/5' splice site interaction.

Class III: All of the slu strains, except slu 8, as well as prp17 and prp18 strains increase cleavage at both of the mutant 5' splice sites; cleavage at the wild-type site is slightly increased or decreased. Curiously, slu4, slu7, prp 17 and prp18 mutants are specifically defective in the second step of splicing (Frank, et al., 1992; Vijayraghavan and Abelson, 1990; Vijayraghavan, et al., 1989); it is not clear why these strains should affect the 5' splice site competition. Notably, slu4 and prp17 (Frank, et al., 1992; Vijayraghavan, et al., 1989), which were independently isolated in

different strain backgrounds, are allelic thus, it is encouraging that both confer the same phenotypes on the competition reporters.

The Future

These data are extremely difficult to interpret because of differences in strain backgrounds. Moreover, most of the mutant strains analyzed in this study were generated by exposure to mutagens and have not been out-crossed. Thus, multiple mutations may be contributing to the observed phenotypes. Thus, it may prove very fruitful to construct a set of isogenic prp strains in the Δcup1 background. Then one could quickly screen through both 5' and 3' splice site competition assays in order to identify proteins which interact with these sites or affect recognition of them. Nonetheless, the results presented here provide interesting hints to the roles of some proteins in 5' splice site choice.

Table 1 Summary of effects conferred by prp and slu strains on competition assays: ratio of cleavage at upstream site relative to downstream site.

strain	WT/WT*	U4A/WT	G5C/WT
wild type	1.9	.37	.19
prp3	0.8	.09	nd
prp4	nd	nd	.15
prp5	1.4	.26	.36
prp7	1.2	.26	.08
prp8	1.9	.36	.36
prp9	1.2	.22	.29
prp11	1.6	.25	.50
prp16-1	3.1	.32	.14
prp17	1.2	.67	.59
prp18	1.2	.48	.59
prp23	1.9	.40	nd
slu1	2.1	.45	.38
slu4	1.2	.71	.59
slu5	2.0	.59	.59
slu6	2.3	.48	.48
slu7	1.9	.53	.26
slu8	3.0	.40	.14
slu9	2.3	nd	.48
slu10	1.8	.32	.31
slu11	1.7	.48	.50
slu12	1.4	.38	.21
U5-U98C	1.2	.17	.07
U5-U98A	1.6	.38	nd
U5-U97CU99C	2.7	.43	.19

*numerator indicates sequence of upstream 5' splice site and denomenator indicates sequence at downstream 5' splice site.

Table 2. Summary of effects conferred by prp and slu strains on competition assays (amount of cleavage in the mutant strains relative to the wild-type).

strain	WT/WT*	U4A/WT	G5C/WT
wild type	WT	WT	WT
prp3	↓↓	↓↓↓	nd
prp4	nd	nd	↓↓
prp5	↓	↓	↑↑
prp7	↓	↓	↓↓↓
prp8	↓	w t	↑↑
prp9	↓	↓↓	↑
prp11	↓	↓	↑↑↑
prp16-1	↑↑	WT	↓↓
prp17	↓	↑↑↑	↑↑↑
prp18	↓	↑	↑↑↑
prp23	WT	WT	nd
slu1	WT	↑	↑↑
slu4	↓	↑↑↑	↑↑↑
slu5	WT	↑↑	↑↑↑
slu6	↑	↑	↑↑↑
slu7	WT	↑↑	↑
slu8	↑↑	WT	↓↓
slu9	↑	nd	↑↑↑
slu10	WT	WT	↑↑
slu11	WT	↑	↑↑↑
slu12	↓	WT	WT
U5-U98C	↓	↓↓	↓↓
U5-U98A	↓	WT	nd
U5-U97CU99C	↑↑	↑	WT

*numerator indicates sequence of upstream 5' splice site and denomenator indicates sequence at downstream 5' splice site.

Table 3 Summary of effects conferred by prp and slu strains on competition assays (In each strain, cleavage at the upstream mutant site has been normalized for cleavage at the wild-type upstream site).

strain	WT/WT*	U4A/WT	G5C/WT
wild type	1.0	.19	.10
prp3	1.0	.11	nd
prp4	nd	nd	.15
prp5	1.0	.19	.26
prp7	1.0	.21	.07
prp8	1.0	.19	.19
prp9	1.0	.18	.24
prp11	1.0	.16	.31
prp16-1	1.0	.10	.05
prp17	1.0	.56	.49
prp18	1.0	.40	.49
prp23	1.0	.21	nd
slu1	1.0	.21	.18
slu4	1.0	.59	.49
slu5	1.0	.30	.30
slu6	1.0	.21	.21
slu7	1.0	.28	.14
slu8	1.0	.13	.05
slu9	1.0	nd	.21
slu10	1.0	.18	.17
slu11	1.0	.28	.29
slu12	1.0	.27	.15
U5-U98C	1.0	.14	.06
U5-U98A	1.0	.24	nd
U5-U97CU99C	1.0	.16	.07

*numerator indicates sequence of upstream 5' splice site and denomenator indicates sequence at downstream 5' splice site.

Table 4 Summary of effects conferred by prp and slu strains on competition assays (amount of cleavage in the mutant strains relative to the wild-type after normalization, see Table3)

strain	U4A/WT	G5C/WT
wild type	WT	WT
prp3	↓↓	nd
prp5	WT	↑
prp7	WT	↓
prp8	WT	↑
prp9	WT	↑
prp11	↓	↑↑
prp16-1	↓↓	↓↓
prp17	↑↑↑	↑↑↑
prp18	↑↑	↑↑↑
prp23	WT	nd
slu1	WT	↑
slu4	↑↑↑	↑↑↑
slu5	↑	↑↑
slu6	WT	↑
slu7	↑	↑
slu8	↓↓	↓↓
slu9	nd	↑
slu10	WT	↑
slu11	↑	↑↑
slu12	↑	↑
U5-U98C	↓	↓↓
U5-U98A	↑	nd
U5-U97CU99C	↓	↓

*numerator indicates sequence of upstream 5' splice site and denomenator indicates sequence at downstream 5' splice site.

REFERENCES

Aebi, M., Hornig, H. and Weissmann, C. (1987) 5' Cleavage site in eucaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU *Cell* **50**: 237-246.

Anderson, J., Wilson, S., Datar, K. and Swanson, M. (1993) NAB2: a yeast nuclear polyadneylated RNA-binding protein essential for cell viability *Molecular and Cellular Biology* **13**: 2730-2741.

Becker, D.M. and Guarente, L. (1991) in *Methods in Enzymology: Guide to yeast genetics and molecular biology* (Guthrie, C. & Fink, G.R., Eds.) pp 182-186, Academic Press, San Diego.

Broach, J. and Volkert, F. (1991) in *The molecular and cellular biology of the yeast Saccharomyces* (J. Broach, J.P.a.E.J., Eds.) pp 297-332, Cold Spring Harbor Press, Cold SPring Harbor.

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

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

Buvoli, M., Biamonti, G., Tsoulfas, P., Bassi, M.T., Ghetti, A., Riva, S. and Morandi, C. (1988) cDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms *Nucleic Acids Res* **16**: 3751-70.

Cech, T. (1993) in *The RNA World* (Gestland, R. & Atkins, J., Eds.) pp 239-270, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Couto, J.R., 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.

Dabeva, M.D., Post-Beittenmiller, M.A. and Warner, J.R. (1986) Autogenous regulaton of splicing of the transcript of a yeast ribosomal protein gene *Proc. Natl. Acad. Sci. USA* **83**: 5854-5857.

Dalbadie-McFarland, G. and Abelson, J. (1990) PRP5: A helicase-like protein required for mRNA splicing in yeast *Proc. Natl. Acad. Sci. USA* **87**: 4236-4240.

Dammel, C. and Noller, H. (1993) A cold-sensitive mutation in 16sS rRNA provides evidence for helical switching in ribose assembly *Genes Dev.* **7**: 660-670.

Dower, W.J., Miller, J.F. and Ragsdale, C.W. (1988) High efficiency transformation of *E. coli* by high voltage electroporation *Nucl. Acids Res.* **16**: 6127-6145.

Engebrecht, J., Voelkel-Meiman, K. and Roeder, G.S. (1991) Meiosis-specific RNA splicing in yeast *Cell* **66**: 1257-1268.

Eperon, L.P., Estibeiro, J.P. and Eperon, I.C. (1986) The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA *Nature* **324**: 280-282.

Fogel, S., Welch, J., Cathala, G. and Karin, M. (1983) Gene amplification in Yeast: CUP1 copy number regulates copper resistance *Curr. Genet.* **7**: 1-9.

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

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

Fu, X.D. (1993) Specific commitment of different pre-mRNAs to splicing by single SR proteins [see comments] *Nature* **365**: 82-5.

Ge, H. and Manley, J.L. (1990b) A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro *Cell* **62**: 25-34.

Gil, A., Sharp, P.A., Jamison, S.F. and Garcia, B.M. (1991) Characterization of cDNAs encoding the polypyrimidine tract-binding protein *Genes Dev* **5**: 1224-36.

Goguel, V., Liao, X., Rymond, B. and Rosbash, M. (1991) U1 snRNP can influence 3'-splice site selection as well as 5'-splice site selection *Genes Dev.* **5**: 1430-1438.

Green, M.R. (1991) Biochemical mechanisms of constitutive and regulated pre-mRNA splicing *Annu. Rev. Cell Biol.* **7**: 559-599.

Guthrie, C. (199) Messenger RNA splicing in yeast: clues to why the spliceosome is a ribonucleoprotein *Science* **253**: 157-163.

Hamer, D., Thiele, D. and Lemontt, J. (1985) Function and autoregulation of yeast copperthionen *Science* **228**: 685-690.

Heinrichs, V., Bach, M., Winkelmann, G. and Luhrmann, R. (1990) U1-specific protein C needed for efficient complex formation of U1 snRNP with a 5' splice site *Science* **247**: 69-72.

Higuchi, R. (1989) in *PCT Technology: Principles and applications for DNA amplification* (Erlich, H.A., Eds.) pp 61-70, Stockton Press, New York.

Hodge, M. and Cumsky, M. (1989) Splicing of a yeast intron containing an unusual 5' junction sequence *Mol. Cell. Biol.* **9**: 2765-2770.

Jacquier, A., Rodriguez, J.R. 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.

Karin, M., Najarian, R., Haslinger, A., Valenzuela, P., Welch, J. and Fogel, S. (1984) Primary structure and transcription of an amplified genetic locus: the CUP1 locus of yeast *Proc. Natl. Acad. Sci.* **81**: 337-341.

Keller, E.B. and Noon, W.A. (1984) Intron splicing: A conserved signal in introns of animal pre-mRNAs *Proc. Natl. Acad. Sci.* **81**: 7417-7420.

Konforti, B.B., Koziolkiewicz, M.J. and Konarska, M.M. (1993) Disruption of base pairing between the 5' splice site and the 5' end of U1 snRNA is required for spliceosome assembly *Cell* **75**: 863-874.

Krainer, A.R., Conway, G.C. and Kozak, D. (1990b) The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites *Cell* **62**: 35-42.

Kumar, A., Sierakowska, H. and Szer, W. (1987) Purification and RNA binding properties of a C-type hnRNP protein from HeLa cells *J Biol Chem* **262**: 17126-37.

Kunkel, T.A., Roberts, J.D. and Zakour, R.A. (1987) Rapid and efficient site-specific mutagenesis without phenotypic selection *Methods Enzymol.* **154**: 367-382.

Larson, G.P., Itakura, K., Ito, H. and Rossi, J. (1983) *Saccharomyces cerevisiae* actin- Escherichia coli LacZ gene fusions: synthetic-oligonucleotide-mediated deletion of the 309 bse pair intervening sequence in the actin gene. *Gene* **22**: 31-39.

Lear, A.L., Eperon, L.P., Wheatley, I.M., Eperon, I.C. and (1991) Hierarchy for 5' splice site preference determined in vivo *JMB* **211**: 103-115.

Legrain, P. and Chapon, C. (1993) Interactions between PRP11 and SPP91 yeast splicing factors and characterization of a PRP9-PRP11-SPP91 complex *Science* **262**: 108-110.

Legrain, P., Chapon, C. and Galisson, F. (1993) Interactions between PRP9 and SPP91 splicing factors identify a protein complex required in prespliceosome assembly *Genes Dev* **7**: 1390-9.

Legrain, P., Seraphin, B. and Rosbash, M. (1988) Early commitment of yeast pre-mRNA to the spliceosome pathway *Mol. Cell. Biol.* **8**: 3755-3760.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L. and Steitz, J.A. (1980) Are snRNPs involved in splicing? *Nature* **283**: 220-224.

Liao, X.C., Colot, H.V., Wang, Y. and Rosbash, M. (1992) Requirements for U2 snRNP addition to yeast pre-mRNA *Nucleic Acids Res* **20**: 4237-45.

Liao, X.C., 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**: 419-28.

Lossom, R. and Lacroute, F. (1979) Interference of nonsense mutations with eucaryotic messeger RNA stability *Proc. Natl. Acad. Sci. USA* **76**: 5134-5137.

Madhani, H. 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.

Mayeda, A. and Krainer, A.R. (1992) Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2 *Cell* **68**: 365-75.

Mayrand, S.H. and Pederson, T. (1990) Crosslinking of hnRNP proteins to pre-mRNA requires U1 and U2 snRNPs *Nucleic Acids Res* **18**: 3307-18.

Moore, M., Query, C. and Sharp, P. (1993) in *The RNA World* (Gesteland, R. & Atkins, J., Eds.) pp 303-58, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Nelson, K. and Green, M. (1988) Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing *Genes Dev.* **2**: 319-329.

Nelson, K., Green, M. and (1991) Mechanism for cryptic splice site activation during pre-mRNA splicing *Proc. Natl. Acad. Sci. USA* **87**: 6253-6257.

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

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

Newman, A.J., R.Lin, S.Cheng and J.Abelson (1985) Molecular consequences of specific intron mutations on yeast mRNA splicing in vivo and in vitro. *Cell* **42**: 335-344.

Ng, R. and Abelson, J. (1980) Isolation and sequence of the gene for actin in *Saccharomyces cerevisiae Proc. Natl. Acad. Sci. USA* **77**: 3912-3916.

Oliphant, A.R., Nussbaum, A. and Struhl, K. (1986) Cloning of random-sequence oligonucleotides *Gene* **44**: 177-183.

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

Parker, R., Siliciano, P. and Guthrie, C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA *Cell* **49**: 220-239.

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, J.G., Mayer, S.A., Tempst, P. and Nadal, G.B. (1991) Characterization and molecular cloning of polypyrimidine tract-binding protein: a component of a complex necessary for pre-mRNA splicing *Genes Dev* **5**: 1237-51.

Reed, R. and Maniatis, T. (1986) A role for exon sequences and splice-site proximity in splice-site selection *Cell* **46**: 681-690.

Rine, J. (1991) in *Methods in Enzymology, Guide to Yeast Genetics and Molecular Biology* (Guthrie, C. & Fink, G.R., Eds.) pp 239-250, Academic Press, Inc., San Diego.

Rio, D. (1993) Splicing of pre-mRNA: mechanism, regulation and role in development *Cur. Opin. Gen. Dev.***3**: 574-584.

Rogers, J. and Wall, R. (1980) A mechanism for RNA splicing *Proc. Natl. Acad. of Sci. USA* **77**: 1877-1881.

Rothstein, R. (1991) in *Methods in Enzymology: Guide to yeast genetics and molecular biology* (Guthrie, C. & Fink, G.R., Eds.) pp 281-301, Academic Press, San Diego.

Ruby, S.W. and Abelson, J. (1988) An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly *Science* **242**: 1028-1035.

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

Rymond, B.C. and Rosbash, M. (1992) in *Molecular and Cellular Biology of the Yeast, Saccharomyces: Gene expression* (Broach, J.R., Pringle, J.R. & Jones, E.W., Eds.) pp CSH Press, Cold Spring Harbor.

Sambrook, J., Fritsch, E.F., Maniatis, T. and *Molecular Cloning: A laboratory manual* (Cold Spring Harbor Laboratory Press, New York, 1989).

Sanger, F.S., Niclen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors *Proc. Nat. Acad. of Sci. USA* **74**: 5463-5467.

Sawa, H. and Abelson, J. (1992) Evidence for a base-pairing interaction between U6 snRNA and the 5' splice site during the splicing reaction in yeast. *Proc. Natl. Acad. Sci. USA* **89**: 11269-11273.

Schena, M., Picard, D. and Yamamoto, K.R. (1990) in *Methods in Enzymology, Guide to Yeast Genetics and Molecular Biology* (Guthrie, C. & Fink, G.R., Eds.) pp 389-398, Academic Press, Inc., San Diego.

Seraphin, B., Kretzner, L. and Rosbash, M. (1988) A U1 snRNA:pre-mRNA base pairing interaction is required early in yeast spliceosome assembly but does not uniquely define the 5' cleavage site. *EMBO* **7**: 2533-2538.

Seraphin, B. and Rosbash, M. (1989a) Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing *Cell* **59**: 349-358.

Seraphin, B. and Rosbash, M. (1989b) Mutational analysis of the U1 small nuclear RNA and pre-mRNA of yeast *Gene* **82**: 145-151.

Sherman, F. (1991) in *Methods in Enzyology:Guide to yeast genetics and molecular biology* (Guthrie, C. & Fink, G.R., Eds.) pp 3-20, Academic Press, San Diego.

Siliciano, P.G. 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.

Steitz, E.S.a.J. (1993) Identification of U5 and U6 small nuclear RNA's as active site components of the spliceosome *Science* in press.

Strathern, J.N. and Higgins, D.R. (1991) in *Methods in Enzymology: Guide to yeast genetics and molecualr biology* (Guthrie, C. & Fink, G.R., Eds.) pp 319-328, Academic Press, San Diego.

Swanson, M.S. and Dreyfuss, G. (1988) RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns *EMBO J* **7**: 3519-29.

Teem, J.L. and Rosbash, M. (1983) Expression of a β-galactosidase gene containing the ribosomal protein 51 intron is sensitive to the *rna2* mutation of yeast *Proc. Natl. Acad. Sci. USA* **80**: 4403-4407.

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.

Wassarman, D.A. and Steitz, J.A. (1992) Psoralen Crosslinking Detects Specific Interactions of U1, U2, U5 and U6 Small

Nuclear RNAs with Precursor Messenger RNA During In Vitro Splicing *Science* **257**: 1918-1925.

Welch, J., Fogel, S., Cathala, G. and Karin, M. (1983) Industrial yeasts display tandem gene iteration at the CUP1 region *Mol. Cell. Biol.* **3**: 1353-1361.

Weser, U., Hartmann, H.-J., Fretzdorff, A. and Strobel, G.-J. (1977) Homologous copper(I)-(thiolate)$_2$-chromophores in yeast copper thionen *Biochim. Biophys.* **493**: 465-477.

Wise, J. (1991) in *Methods in enzyology: Guide to yeast genetics and molecular biology* (Guthrie, C. & Fink, G.R., Eds.) pp 405-414, Academic Press, San Diego.

Woolford, J., J.L. (1989) Nuclear pre-mRNA splicing in yeast *Yeast* **5**: 439-457.

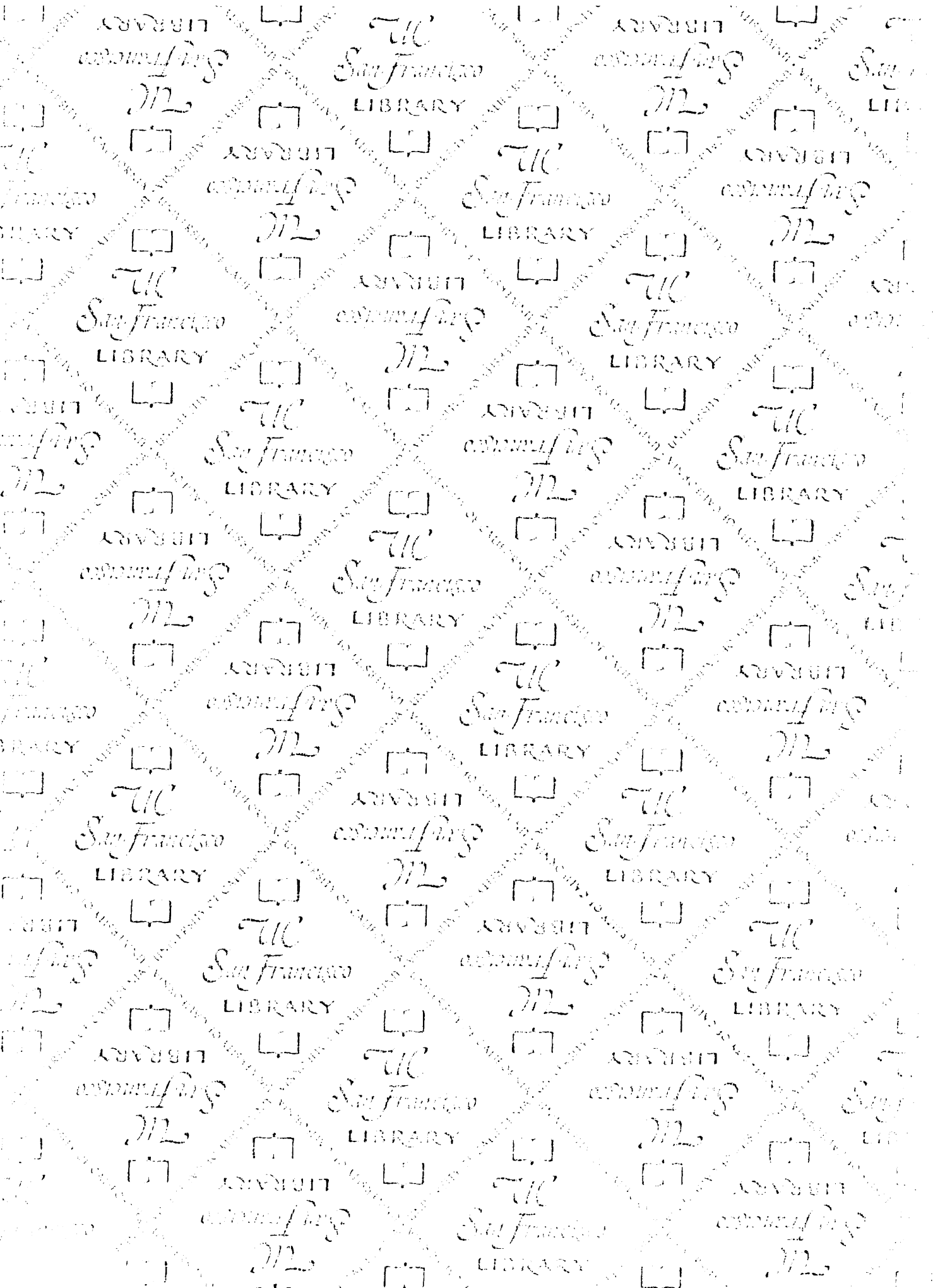
Wu, J. and Manley, J.L. (1989) Mammalian pre-mRNA branch site selection by U2 snRNP involves base-pairing *Genes Dev.* **3**: 1553-1561.

Wyatt, J., Sontheimer, E. and Steitz, J. (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.

Zahler, A.M., Neugebauer, K.M., Lane, W.S. and Roth, M.B. (1993) Distinct functions of SR proteins in alternative pre-mRNA splicing *Science* **260**: 219-22.

Zhuang, Y. and Weiner, A.M. (1986) A compensatory base change in U1 snRNA suppresses a 5' splice site mutation *Cell* **46**: 827-835.

Zhuang, Y. and Weiner, A.M. (1989) A compensatory base change in human U2 snRNA can suppress a branch site mutation *Genes & Dev.* **3**: 1545-1552.


San Francisco
LIBRARY