Site-specific Chemical Modifcation and Crosslinking Studies of U6 snRNA in the Yeast Spliceosome

Thesis by

Chang Hee Kim

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1999

(Defended September 17, 1998)

Acknowledgements

Above all, I would like to thank our Lord Jesus Christ for the opportunity, health, intellect, wisdom and the persistence to be able to complete my Ph.D, and for giving me purpose and fulfillment in my life. I would also like to acknowledge all the loving members of the Logos church youth group for their prayers which made this possible.

Great thanks are due to Professor John Abelson for providing the resources, motivation and the intellectual freedom where excellent research could be conducted. I would like to thank him especially for the confidence he had in my abilities to come through during difficult times

Thanks are due to the many post-doctoral fellows of the Abelson lab over the years for their support scientifically and otherwise. I am particularly indebted to Drs. Jeff Sampson, Peggy Saks, Tadeusz Marciniec, Dan Ryan, Andrea Ghetti, Hitoshi Sawa, John Wagner, Chris O'Day, Debra Wiest, Scott Stevens and David McPheeters.

I can't leave out my fellow Entruvidorian counterpart Chris Trotta for his friendship and support over the years as well as the "replicase" and other scientific counsel and camaraderie.

Thanks to Hoiwon, Hyejin, Keith Kuwata, Johan Chu for their lasting friendship and the good times.

Thanks to Esther for her love and for our future together.

Finally, I would like to thank my mother and my father for their care and love and most of all prayer.

Abstract

Pre-mRNA splicing is an essential step in eukaryotic gene expression. Most eukaryotic genes are interrupted by sequences, called introns, that do not code for sequences in the final protein gene product. Pre-mRNA splicing is the process by which introns are removed and the coding sequences, called exons, are joined together following transcription of a gene into RNA. The spliced mRNA subsequently acts as the coding template for protein synthesis. Pre-mRNA splicing occurs in a large ribonucleoprotein particle called the spliceosome which includes five essential small nuclear RNAs (snRNAs) and more than 50 different proteins. It is thought that the active site of the spliceosome consists of U2, U6 and U5 snRNAs. In order to probe the structure of the catalytic site, we have done an extensive site-specific chemical modification and crosslinking studies of U6 RNA in the yeast spliceosome. From an extensive screen of site-specific 2'-deoxy modifications, we have found four that block the first step of splicing, yet are able to assemble spliceosome complexes which precede the reactive spliceosome, suggesting that the 2'-hydroxyls at these nucleotides may be required for catalysis. In the crosslinking experiments,we found crosslinks between 4-thioU placed in the central conserved sequences in U6 with the 5' splice site of the pre-mRNA, which participates in the first chemical step of splicing. We were able to determine the order of this crosslink by doing the experiment in yeast mutant spliceosomes which were blocked in well-defined stages of splicing. We also found that a conserved nucleotide in the 3' stem-loop of U6 crosslinks to a nucleotide in a region of the actin pre-mRNA intron which is required for its splicing in vitro. We also found crosslinks to snRNAs -- U2 and U4. Certain positions in U6 can crosslink first to U4 snRNA, then subsequently to the pre-mRNA, consistent with the notion that the spliceosome is a dynamic machine.

Table of Contents

Chapter I - Review: Pre-mRNA Splicing. (John Abelson and Chang Hee Kim, published in Frontiere della Biologica ed. Andrea Turchi. Roma, Italy, 1998)

Abstract

Most eukaryotic genes are interrupted by sequences which ultimately do not code for any amino acid sequence. These non-coding sequences (introns) are removed and the coding sequences (exons) joined together post-transcriptionally by a process called RNA splicing. There are four different types of splicing: pre-tRNA splicing (catalyzed by protein enzymes), Group I and Group II self-splicing (catalytic RNAs), and pre-mRNA splicing (which occurs in a huge ribonucleoprotein particle called the spliceosome). Five essential small nuclear RNAs (snRNAs) assemble on the pre-mRNA substrate in a step-wise manner in conjunction with many proteins. It is thought that the active site of the spliceosome is composed of RNA (the conserved snRNAs) because pre-mRNA splicing follows the same chemical mechanism as Group II self-splicing catalytic RNA. Some proteins in the spliceosome are thought to play a structural (e.g. snRNPs) or regulatory role (e.g. SR proteins), while others are thought to be ATPases or RNA helicases (DEAD box proteins). A new type of spliceosome, which is used for splicing AT-AC introns also shows similarities in U snRNP content and structure to the GU-AG spliceosomes.

Introduction

Following the discovery in 1953 of the structure of DNA and its potential for the storage of genetic information, the idea that the nucleotide sequence of a gene is colinear with its polypeptide product became a central tenet of Biology. Work with bacteria and

phage in the 60s and 70s provided ample proof of this notion so it was a shock to biologists that the first comparisons of eukaryotic mRNA sequences with that of their genes demonstrated that eukaryotic genes are often interrupted by sequences which do not code for the polypeptide product of the gene. Thus the gene for the β subunit of hemoglobin has two intervening sequences or introns which interrupt the colinearity of the gene and its product. These sequences are not found in β hemoglobin mRNA and must be removed from a precursor mRNA following transcription of the gene. The introns are removed by an RNA processing event called splicing (Figure 1). The entire gene, including the introns is transcribed and the introns are removed from the pre-mRNA in a precise splicing reaction in which each intron sequence is released by two endonucleolytic cuts and the ends are rejoined to form the mature RNA.

As cloning and DNA sequencing techniques improved it was soon realized that some genes are far more complicated than the β hemoglobin gene (Figure 2). Vertebrate collagen genes, for example contain more than 50 introns raising the requirement for extreme accuracy of the splicing process (for example a one nucleotide mistake in one of the 50 splicing events would create a frameshift in the mRNA product and translation would result in a missense product). In some genes, particularly some of the homeotic Drosophila

Figure 1. Pre-mRNA Splicing is an essential post-transcriptional step in gene expression. Exons are denoted by blue bars, and intron by thin black line.

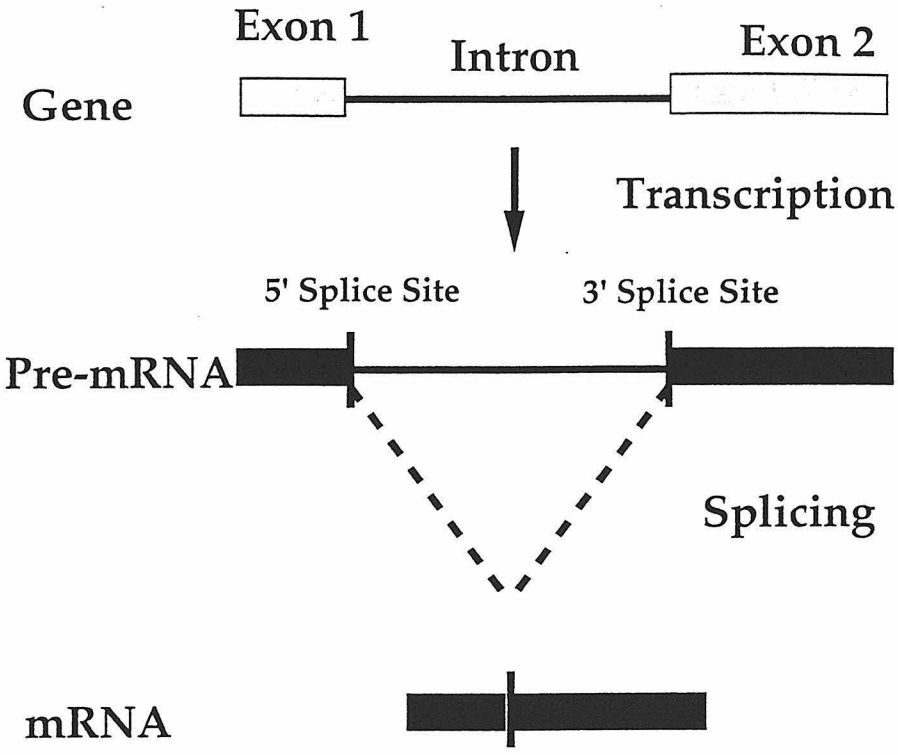
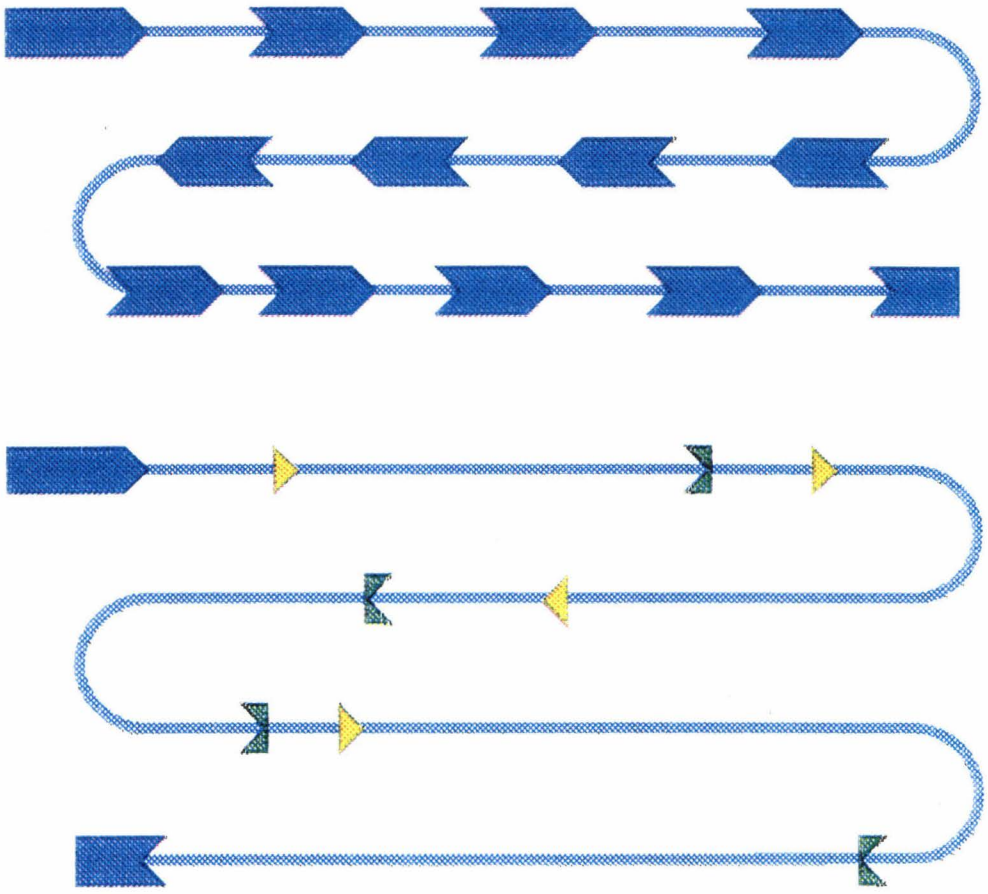

Exon 1
Intron
Exon 2
Gene
Transcription
5' Splice Site
3' Splice Site
Pre-mRNA
Splicing
mRNA

Figure 2. Intron/exon topology. In some genes, the intron can be very large. In others, there are a large number of introns.



genes such as bithorax, there are very large introns, as long as one million bases, raising the problem of how the splice sites are recognized in this sea of RNA information. In many cases the splicing of complex genes is regulated. In different tissues alternative splicing patterns are followed for a particular pre-mRNA making possible the controlled synthesis of more than one protein for each gene.

When introns were discovered in 1977 it seemed likely that as in protein synthesis or transcription there would be basically one mechanism of splicing. As it turned out there are at least four mechanisms. In each of the four mechanisms the structure of the precursor RNA is different. For the splicing of introns from pre-mRNAs there is no conserved structure of the intron. Instead the intron boundaries are designated by minimal consensus sequences as seen in Figure 3.

Some tRNA genes in both eukaryotes and in archaebacteria contain small introns which usually interrupt the anticodon loop. In this case there are no conserved sequences at the splice junctions and the intron is recognized by its position in the tertiary structure of the pre-tRNA.

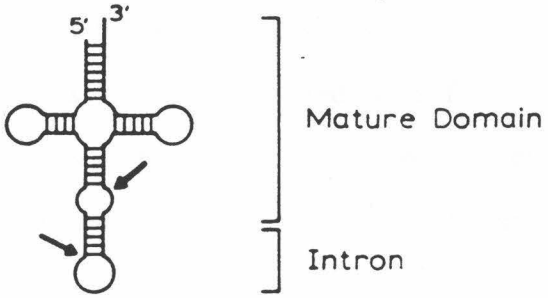
In one of the most exciting and unexpected discoveries in recent times it was found that there are two different classes of self-splicing introns. These catalytic RNAs have the ability to splice themselves from a transcript. In this case it is the secondary and tertiary structure of the intron which is important and the two different self-splicing intron groups, class I and class II, have distinct secondary and tertiary structures.

The chemical mechanisms for each of the four types of splicing is known. In tRNA splicing in eukaryotes, the reaction is catalyzed by three enzymes: an endonuclease, a ligase and a phosphotransferase (Figure 4). The endonuclease cleaves pre- tRNA at the 5' and 3' splice sites to release an intron. The products of the endonuclease reaction are two tRNA half-molecules bearing a 2'-3' cyclic phosphate and 5'-OH termini (Peebles *et al.*, 1983) The tRNA ligase, through a complex series of ATP- and GTP- dependent reactions, covalently joins the half-molecules (Westaway *et al.*, 1993; Westaway *et al.*, 1988). After

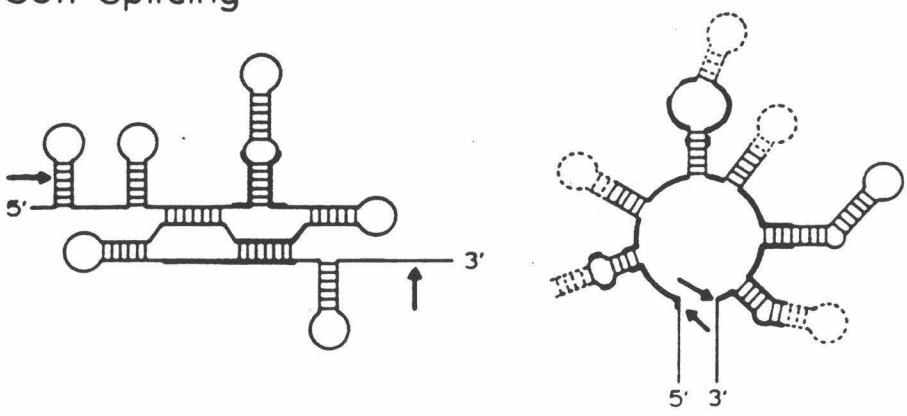
Figure 3. Four kinds of introns, four kinds of splicing. depicted are the consensus sequences in yeast introns, the secondary structure of yeast tRNA precursors, and consensus secondary structures of group I and group II self-splicing introns. Splice junctions are indicated by arrows.

■ Nuclear mRNA

■ tRNA



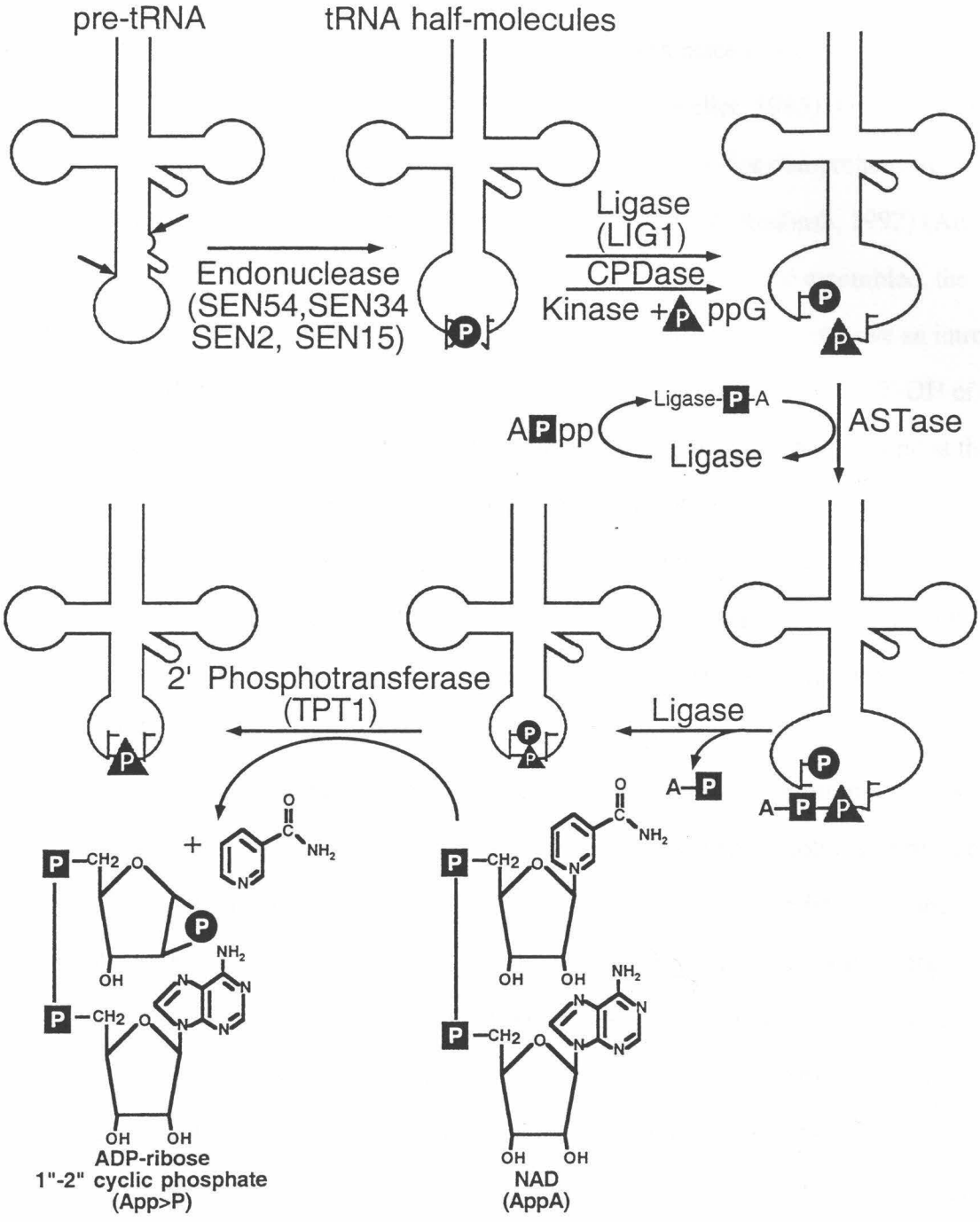
■ Self-Splicing



Group I

Group II

Figure 4. Mechanism for tRNA Splicing in yeast, as described in text. Splice sites are indicated by small arrows, and the intron is shown in bold. The phosphate from exogenous GTP is indicated by a filled-in triangle and the endogenous phosphate by a filled circle. CDPase = cyclic phosphodiesterase; ASTase = adenylyl synthetase; 2'-PTase = 2'-phosphatase.


pre-tRNA
tRNA half-molecules
Endonuclease
(SEN54,SEN34
SEN2, SEN15)
Ligase
(LIG1)
CPDase
Kinase +
ppG
ASTase
Ligase-P-A
Ligase
2' Phosphotransferase
(TPT1)
Ligase
ADP-ribose
1"-2" cyclic phosphate
(App>P)
NAD
(AppA)

ligation, a 2' phosphate remains at the splice junction and is removed by a phosphotransferase. This enzyme catalyzes transfer of the 2' phosphate to the nicotinamide 2'-OH of NAD (Culver *et al.*, 1993).

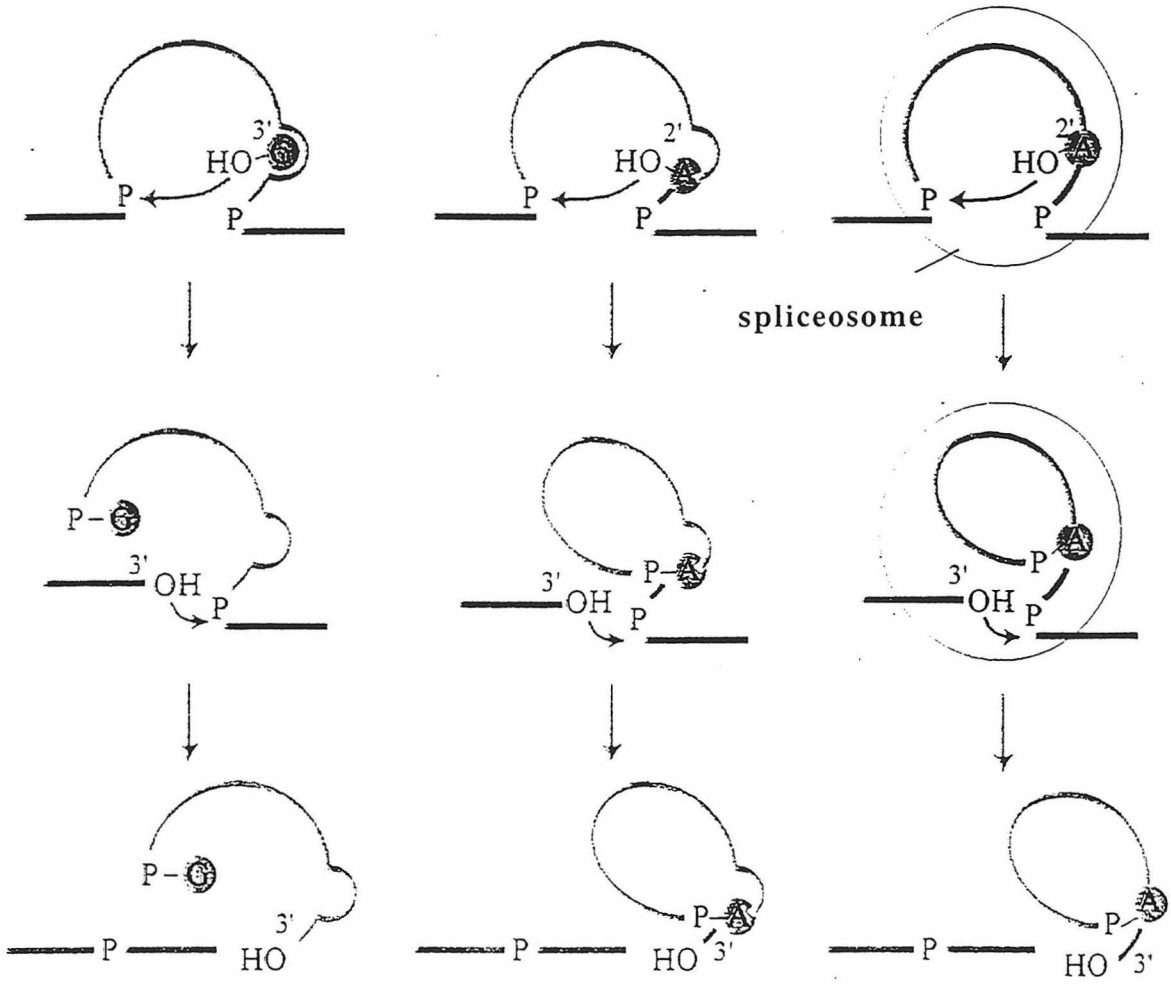
The removal of introns from nuclear pre-mRNA takes place in a complex structure, the spliceosome ((Brody & Abelson, 1985) (Frendewey & Keller, 1985) (Grabowski *et al.*, 1985)), which is assembled de novo for each splicing event (for comprehensive reviews see: (Green, 1991; Ruby & Abelson, 1991) (Rymond & Rosbash, 1992) (Ares & Weiser, 1995; Madhani & Guthrie, 1994a; Moore *et al.*, 1993)). Once assembled, the spliceosome catalyzes two sequential transesterification reactions which remove an intron from the pre-mRNA and ligate the exons (Figure 5). In the first reaction, the 2'-OH of a particular adenosine residue near the 3' splice site attacks the phosphodiester bond at the 5' splice site to produce two intermediates: exon 1 and lariat intron-exon 2, a "branched" or "lariat" structure in which the 5' end of the intron is connected to the branch point adenosine via a 2'-5' phosphodiester bond. In the second reaction, the terminal 3'-OH of exon 1 attacks the phosphodiester bond at the 3' splice site, resulting in formation of the spliced mRNA product and the released intron in lariat form.

Figure 5 compares the mechanism of splicing in the self splicing introns to that of pre-mRNA splicing. In both cases the reaction consists of two phosphotransfer reactions. In Group I splicing, the reaction requires a guanosine co-factor. In the first phospho-transfer reaction, the 3'-OH of the guanosine attacks the phosphodiester bond at the 5'-splice site transferring the G to the 5' end of the intron in a 3'-5' phosphodiester bond and giving the intermediates in the reaction, the 5' exon and the intron-3' exon containing the guanosine co-factor at the 5'-end. In the second reaction the 3'-OH of the 5' exon attacks the phosphodiester bond at the 3' splice junction giving the products of the reaction, the spliced RNA and a G-terminated intron. In a series of elegant experiments, Cech and his colleagues have shown that the intron itself continues to be catalytically active and can carry out a number of other phosphotransfer reactions.

Figure 5. Mechanisms of Self-splicing (Group I and Group II introns) and Pre-mRNA Splicing. The exons being joined are shown in blue and yellow, and the attacking unit in green. The catalytic site is formed by the intron itself in group I and group II splicing. In pre-mRNA splicing, the snRNAs are thought to compose the active site.

Self-splicing introns

Spliceosome-catalyzed splicing of nuclear mRNA



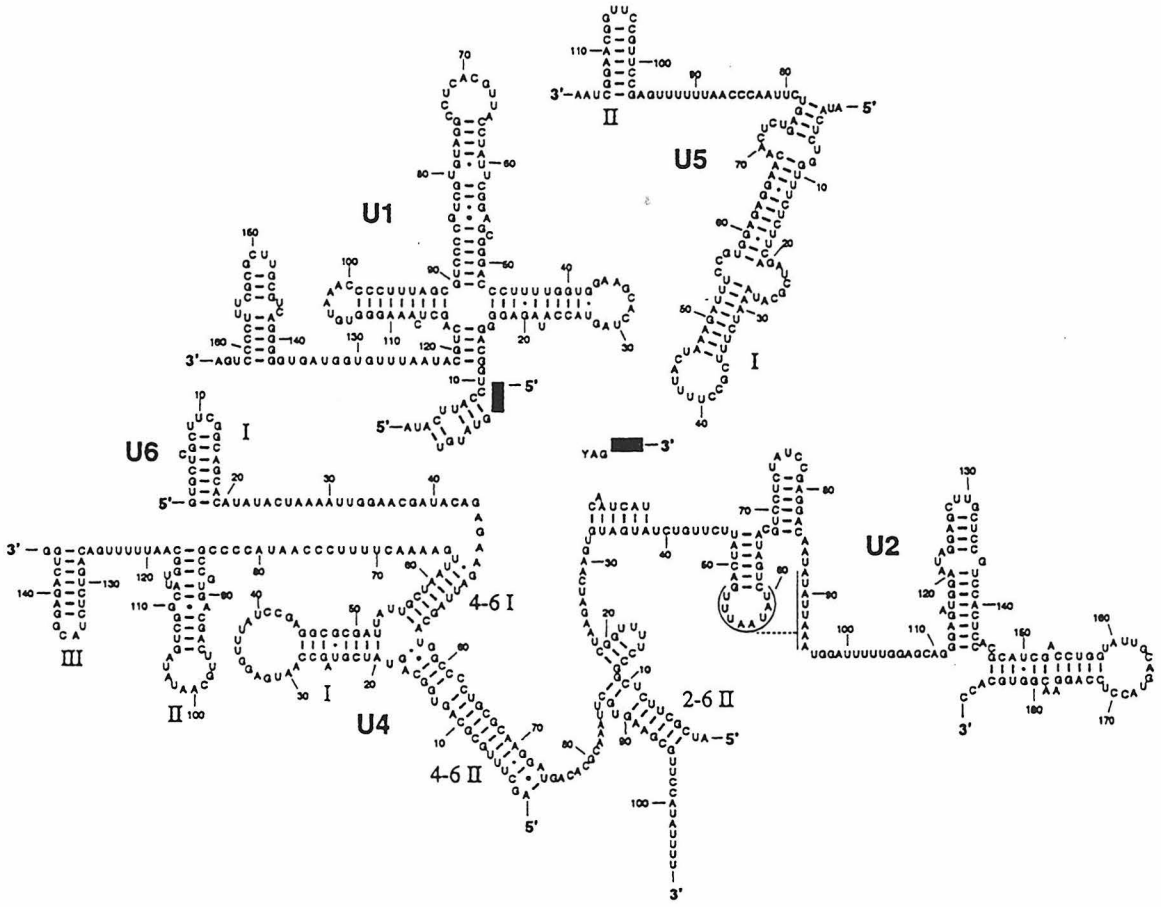
In a discovery which galvanized the RNA processing field, Peebles (Peebles *et al.*, 1986) and Van der Veen (Veen *et al.*, 1986) showed that the group II self splicing introns are spliced by an identical pathway to that of pre-mRNA splicing, leading to the principal tenets of this field of research: that pre-mRNA splicing is an RNA catalyzed reaction and that the two processes share a common evolutionary ancestor. The two reactions, though mechanistically similar, have quite different requirements. In general, the self splicing introns have no required factors, though in the case of some group II introns a protein is required, probably to stabilize the tertiary structure of the intron. By contrast pre-mRNA splicing requires five additional RNA molecules, the snRNAs, more than 100 proteins and hydrolyzable ATP. What are the roles of the trans-acting RNA and protein molecules in this reaction? Why is ATP required when the chemistry, if it is similar to the Group II splicing mechanism, does not require it. The present hypothesis, being tested in many laboratories, is that the snRNAs, interacting with the pre-mRNA substrate, form an RNA structure which catalyzes the reaction just as the intron does in group II splicing but they must be assembled step-wise into a catalytic structure and this process, splicesome assembly, is required to construct a catalytically active RNA structure around each of the many diverse introns that are spliced in the cell. The hydrolysis of ATP appears to be required at several stages of the assembly process. Thus two principal problems need to be solved in pre-mRNA splicing: 1. What is the catalytically active structure for pre-mRNA splicing, i.e. what is the active site of the enzyme? and 2. How is this enzyme assembled correctly on each of the diverse substrates it must process?

The Spliceosome

Though the biochemistry of pre-mRNA splicing has now been studied in many eukaryotic organisms including yeast, nematodes, flies, plants, frogs and mammals, most of the detailed information on the components of the spliceosome come from studies of the yeast, *Saccharomyces cerevisiae,* and of humans. The two studies complement each other. In the human system, active nuclear extracts from tissue culture cells are prepared and fractionated to elucidate the splicing components and their functions. In yeast a combination of genetics and biochemistry has resulted in the identification and functional characterization of many spliceosomal components. Fortunately the information from the two systems is converging and there are now many examples of components first discovered in yeast whose counterpart in mammals is now known and vice versa. Clearly the mechanism of pre-mRNA splicing is conserved over the entire eukaryotic kingdom though there are some interesting differences being discovered between yeast and mammals.

Five RNAs, U1, U2, U4, U5 and U6 snRNAs, are required for pre-mRNA splicing (Figure 6). U1, U2, U4 and U5 RNAs have a distinct hyper methylated cap structure, m3G, unique to snRNAs. U6 has a γ-phosphoryl methyl ester at its 5' end. In addition each of the U RNAs contains a variety of other modified bases such as pseudouridine and various methylated nucleosides. The snRNAs are found in the cell as ribonucleoprotein particles, RNPs, and thus the name for these particles, snRNPs. Four separate snRNPs function in splicing U1, U2, U5 and U4/U6. In the latter case two RNA molecules, U4 and U6 held together by 21 base-pairs are found in the same particle. Each of the snRNPs has a set of common core proteins, the Sm proteins, which bind to the sequence $AU_{4-6}G$ found in U1, U2, U4 and U5 RNAs. In addition each snRNP has a set of unique proteins. An intermediate in the formation of the spliceosome, the U4/U6-U5 tri-snRNP has been

Figure 6. The small nuclear RNAs (from (Ares & Weiser, 1995)). The sequences of the human spliceosomal snRNAs are folded to show interaction established during spliceosome assembly. Extensive interaction between U4 and U6, as well as between the 5' end of U2 and the 3' end of U6, is shown. The intron sequences connecting the 5' splice site to the branchpoint and the branchpoint to the 3' splice site are not shown. Exon 1, 5' black bar; exon 2, 3' black bar.


U1
U2
U4
U5
U6
I
II
III
4-6 I
4-6 II
2-6 II
5'
3'

purified from Hela cell extracts ((Behrens & Lührmann, 1991) and it contains an additional 5 proteins specific to that structure.

In addition to the proteins found stably associated with snRNPs in extracts, there are a number of proteins that are required for splicing but which are not found in snRNPs. In cases where we know something about the function of these proteins, they appear to mediate specific steps in spliceosome assembly-a topic we will explore in depth later.

Many of the spliceosomal proteins have first been identified in yeast via genetic screens or selections. One of the most productive approaches has been the screening of banks of temperature sensitive yeast mutants for defects in pre-mRNA splicing. The first such bank was constructed more than 30 years ago by Lee Hartwell. Hartwell and McLauglin screened this bank for defects in RNA synthesis-rRNA synthesis it turned out-because most RNA in yeast is rRNA. This screen yielded 11 complementation groups, RNA1-RNA11. Of these RNA2-11 all turned out to be mutants in pre-mRNA splicing. It took some time to realize this since at the time mRNA splicing had not yet been discovered. It turns out that very few yeast genes have introns but almost all of the ribosomal protein genes contain an intron. Thus a deficiency in splicing markedly affects the synthesis of ribosomal proteins and in turn the stability of rRNA. A further screen of a bank of 1000 ts mutants yielded more genes and since then the number has continued to grow (Vijayraghavan *et al.*, 1989). These genes are now called PRP (for pre RNA processing). A summary of the PRP genes is given in Table 1. Not only has this genetic approach been valuable in defining proteins required for splicing but the mutants have been important in defining the steps of spliceosome assembly. For example, when prp2-1 mutant extracts are prepared they are active at the permissive temperature (16-20^{o}C). But inactive after pre-incubation at 37-42^{o}C. The inactive extracts assemble spliceosomes but are blocked at a stage just prior to the splicing reaction. The inactive extract can be complemented by addition of active Prp protein. This experiment can be performed with most of the prp mutants and has been an invaluable aid to dissecting the pathway of spliceosome assembly.

Table 1. Prp proteins in yeast (adapted from (Beggs, 1993)).

Gene	Step	Association/Activities	Sequence motifs
PRP2	1	Spliceosome	DEAH; zinc-finger
PRP3	1	U4/U6 snRNP	
PRP4	1	U4/U6 snRNP	β-subunit of G protein
PRP5	1	U2 addition	DEAD
PRP6	1	U4/U6.U5	TPR/PW, zinc-finger
PRP7	1		
PRP8	1	U5	Proline-rich, acidic N-end
PRP9	1	U2 addition	zinc finger-like
PRP11	1	U2 addition	zinc finger-like
PRP16	2	Spliceosome	DEAH
PRP17/ SLU4	2		β-subunit of G protein
PRP18	2	U5 snRNP	
PRP19	1	Spliceosome	
PRP21/ SPP91	1	U2 addition	surp of *su(w^a)**
PRP22		mRNA release	
PRP24	1	U6snRNP	RNA binding motifs
PRP25	1		
PRP26	Intron Turnover	Debranching activity	
PRP27	Intron Turnover		
PRP28	1		DEAD
PRP29-34		All accumulate pre-mRNA	
PRP38	1		Acidic, serine-rich C-end
PRP39	1	U1 snRNP	
PRP40			
SPP2	1		
SLU1	1		
SLU2	1		
SLU7	2		Zinc knuckle-like
SNP1		U1 snRNP	RNA binding motif
MUD1		U1 snRNP	RNA binding motif
MUD2		commitment complex	
SMD1	1	core snRNP	
BBP		pre-mRNA branchpoint	RNA binding motif
CUS 1			
HSH 49			
BRR1-BRR5			

Spliceosome assembly

Spliceosome assembly, diagrammed in Figure 7A begins with the interaction of the U1 snRNP with the pre-mRNA. The sequence at the 5' end of U1 RNA is complementary to conserved intron nucleotides at the 5' splice site, and base-pairing of these regions is essential. The U1-pre-mRNA complex is called the "commitment complex" because once formed, the pre-mRNA is committed to go through the splicing reaction. It is the only step in spliceosome assembly which does not require ATP hydrolysis.

In the next step, the U2 snRNP joins the commitment complex to form a stable complex called the "pre-spliceosome". In this step a sequence in U2 near its 5' end base-pairs with the conserved intron branch point sequence near the 3' splice site. Formation of the pre-spliceosome requires at least four of the Prp proteins, Prp 5, 9, 11 and 21. In mammalian cell extracts this step requires the action of a well studied protein called U2AF which appears to facilitate the pairing between U2 and the branch point sequence (Valcarel *et al.*, 1996).

U4 and U6 are found in a single snRNP containing a U4/U6 RNA duplex involving 21 base pairs in two helices (Guthrie & Patterson, 1988). The U4/U6 snRNP is joined with the U5 snRNP in a triple snRNP which joins the spliceosome in the next step of assembly.

In this step, a crucial and quite remarkable event occurs. U4 either leaves the spliceosome or its interaction is greatly destabilized as both U4/U6 RNA helices dissociate. The presence of U4 is not necessary for either of the catalytic reactions (Yean & Lin, 1991). Presumably at this point the catalytic center of the spliceosome is nearly complete. (Figure 7B) However, one further ATP-dependent step, mediated by yeast protein Prp2, is required to activate the first step of splicing. In this step a marked conformational change occurs in the spliceosome, but the nature of this change has not yet been defined (Kim & Lin, 1993).

Figure 7A. Spliceosome assembly pathway based on results from yeast and mammalian systems. The abbreviations CC, B, A2-1, A1, A2-2, and A2-3 represent ribonucleoprotein complexes (as defined in yeast) that have been characterized biochemically or genetically. In each distinct complex, many protein factors bind to the complex, but have been omitted for simplicity or lack of data; only the Prp proteins, which are required for the progression from one complex to another, are shown where applicable. The five small nuclear ribonucleoproteins (snRNPs) -- U1, U2, U4, U5, and U6 -- are enclosed and highlighted in circles.

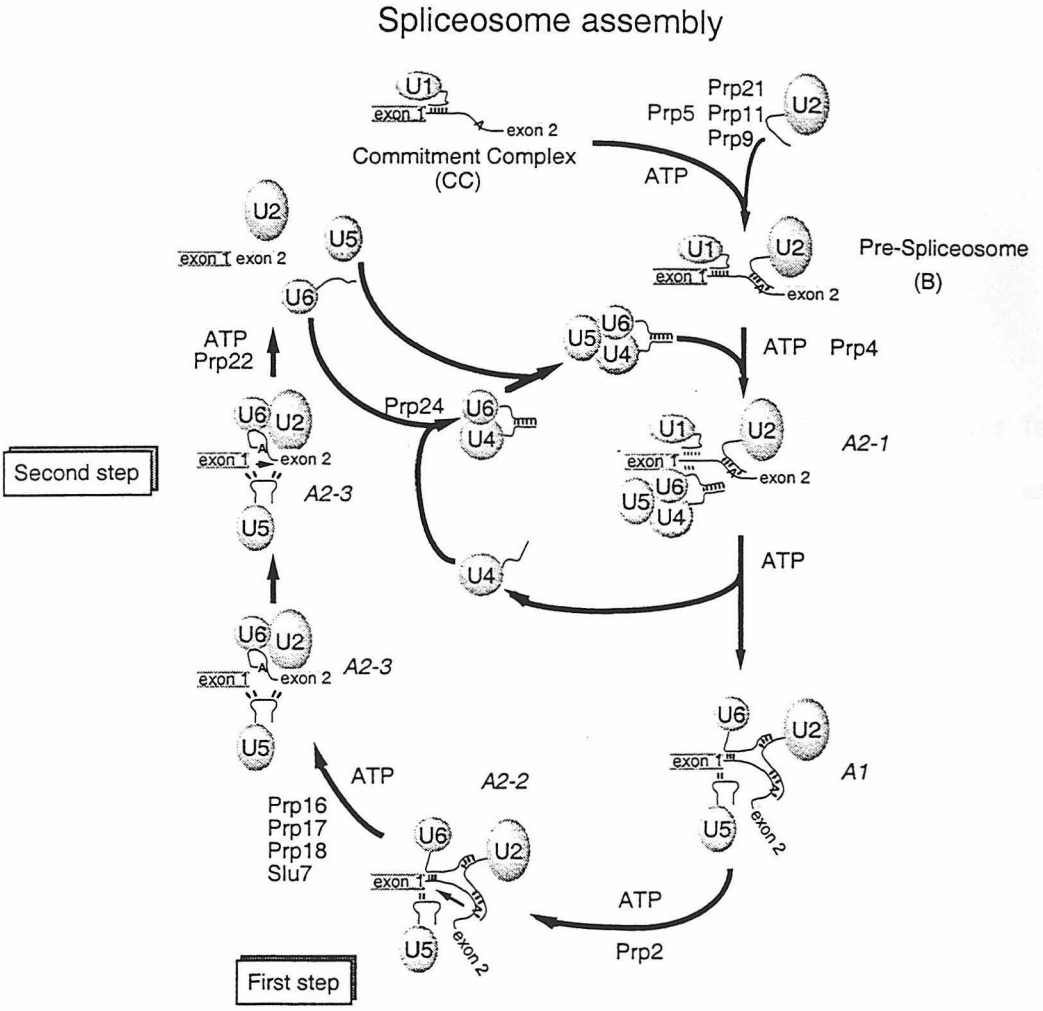
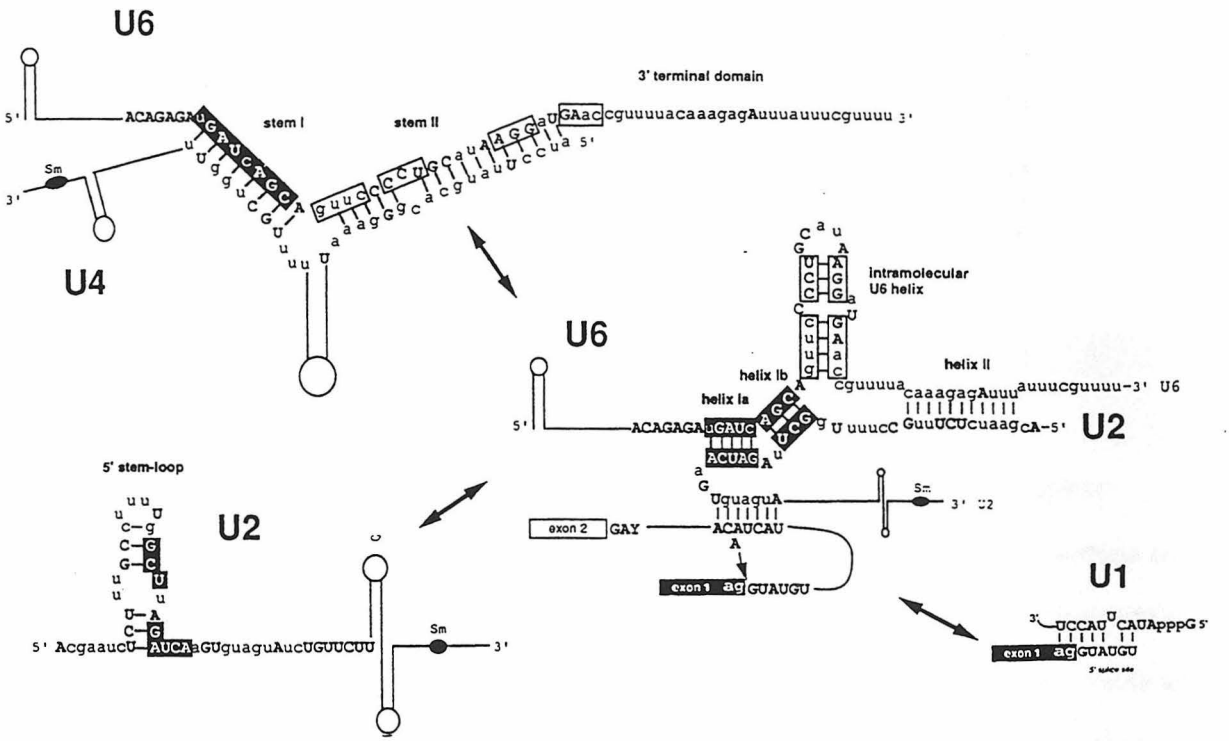

Spliceosome assembly
U1
exon 1
exon 2
Commitment Complex
(CC)
Prp21
Prp5 Prp11
Prp9
U2
ATP
Pre-Spliceosome
(B)
U5
U6
U4
ATP Prp4
ATP
Prp22
Prp24
A2-1
A2-3
Second step
ATP
A2-3
A1
ATP
A2-2
Prp16
Prp17
Prp18
Slu7
ATP
Prp2
First step

Figure 7B. RNA-RNA interactions in the fully assembled spliceosome for the first chemical step. The intermolecular pairing between U6 and U2 and with the pre-mRNA are thought to position the 2'-OH of the branchpoint A close to the 5' splice site.


U6
3' terminal domain
stem I
stem II
Sm
U4
intramolecular
U6 helix
helix Ib
helix Ia
helix II
U6
U2
5' stem-loop
U2
exon 2
GAY
ACAUCAU
exon 1
ag
GUAUGU
U1
Sm

Activation of the spliceosome for the second step of splicing requires yet another ATP- dependent step and presumably another conformational change. At least four yeast proteins are required for this step, Prp 16, 17, 18 and Slu7 (reviewed by Umen and Guthrie 1995b (Umen & Guthrie, 1995a)). Thus it appears that at least a subtle change in the active site occurs between the two steps, an assumption supported by the fact that the stereochemical requirements for the phosphodiester transesterification reaction differ for the first and second step (Moore & Sharp, 1993).

Following the completion of the splicing reaction at least one more ATP dependent step is required for the release of the mRNA product from the spliceosome. This step is mediated by Prp22. The intron remains with the spliceosome and in a series of important, but less well studied steps, the intron RNA is degraded and the snRNPs recycled. Among these reactions is the re-pairing of U6 with U4 and the re-formation of the triple snRNP.

RNA Interactions in Spliceosome Assembly: The Stepwise Assembly of the Active Center.

In spliceosome assembly, the first task is to identify the intron and the splice junctions and then to assembly the spliceosome in such a way that the splice junctions are brought together in the active center of the spliceosome enabling the two phospho-transfer reactions to proceed. Of the five snRNAs, two-U1 and U4, function only in assembly and are either absent or unstably associated with the functional spliceosome. (Most likely U1 and U4 do not really dissociate in vivo but in vitro their interactions with the spliceosome are destabilized. Most important they are not required for the chemistry of splicing.) Abundant data now support obligatory roles for U2, U5 and U6 RNAs in both assembly and catalysis.

The signals which identify portions of a transcript as intron are minimal. In yeast they are the 5' splice site sequence GUAUGU, the branch point sequence UACUAAC and the 3' splice site sequence, PyAG. In mammals these are the consensus sequences but they

are less conserved, especially in the branch point sequence and in addition a poly pyrimidine stretch in the intron near the 3' splice site, usually present as well in yeast introns plays a more important role.

The initial recognition of the intron is by the U1 snRNP which base-pairs with the 5'-splice site via nucleotides near its 5' end. There is a very interesting difference between yeast and mammalian introns as to *which* 5' splice site is recognized. In mammals, whose genes often have many introns in which the exons are small and introns large, the 5' splice site in the down stream intron is recognized by the U1 snRNP and in subsequent assembly steps the U2 and U5-U4/U6 snRNPs are recruited to the upstream 5' splice site. This mode of recognition is called exon definition. (Berget, 1995). Following assembly of this complex across the down stream exon, there must be a rearrangement of the spliceosome across the intron to allow the correct splicing to take place. By contrast in yeast whose genes in general have small single introns, recognition is by intron definition. Though U1 pairing is in general required for definition of the 5' splice site, precise identification of where the cut will be made can also be mediated by U5 interaction with exon bases adjacent to the 5' splice site.

Though the U1 snRNA only pairs with the 5' splice site, there is considerable evidence that even at this early stage the branch point region is also recognized (Figure 7B). This is apparently not via a base-pairing interaction between U1 RNA and sequences near the 3' splice site but rather is mediated by protein-protein interactions.

In the next step, the U2 snRNP joins the commitment complex to form a stable complex called the "pre-spliceosome". In this step a sequence in U2 near its 5' end base-pairs with the conserved intron branch point sequence near the 3' splice site. In this pairing, the penultimate A in the sequence UACUAAC, destined to form the branch point, is bulged out from the helix.

In extracts, the U5 snRNP and the U4/U6 snRNP join to form a triple snRNP and it is apparently as the triple snRNP that these snRNPs join the spliceosome. With this event

comes a number of crucial RNA-RNA interactions-in some cases new-in some cases a replacement of pre-existing interactions. The precise order in which these interactions take place is not known. Many of the steps following the formation of the pre-spliceosome involve changes in preexisting RNA structure-often the exchange of one set of helical interactions for another. The first of these involves the replacement of the U1-5' splice site base pairing for U6-5'-splice site interactions. The 5' end of the highly conserved ACAGAGA sequence of U6 base pairs with the 5' splice site, replacing U1/intron interactions and destabilizing interactions between U1 and the spliceosome. The minimal base-pairing interaction, supported by yeast genetic experiment, is between U6 and the 5' splice site involves the conserved ACA sequence of U6 with the conserved UGU of the intron (Lesser & Guthrie, 1993). However recent crosslinking experiments demonstrate an antiparallel interaction between U6 and the intron which extends over nearly a complete helical turn (Kim & Lin, 1993). In most introns, however, this interaction will not involve perfect base-pairing.

In addition to the U6 5'-splice site interactions, two nucleotide bases in the conserved central loop of U5 snRNA interact with the last two bases of exon 1 (Newman & Norman, 1991; Newman & Norman, 1992) (Sontheimer & Steitz, 1993). While the locale of the 5' splice site is defined by the earlier U1-5'-splice site interactions, the U6 and U5 interactions with the 5' splice site, occuring later in spliceosome assembly, both function in defining the precise phosphodiester bond to be attacked in the first step of splicing. As in later steps we perceive a redundancy in the mechanism of splice site identification, doubtless necessary to achieve the needed accuracy in the reaction.

In the next step of assembly a crucial and quite remarkable event occurs. U4 either leaves the spliceosome or its interaction is greatly destabilized as both U4/U6 RNA helices dissociate. The presence of U4 is not necessary for either of the catalytic reactions (Yean & Lin, 1991). Dissociation of the U4/U6 helices allows U6 to form a U2/U6 helix and an intramolecular stem-loop near the 3' end of U6 (Madhani & Guthrie, 1992) (Fortner *et al.*,

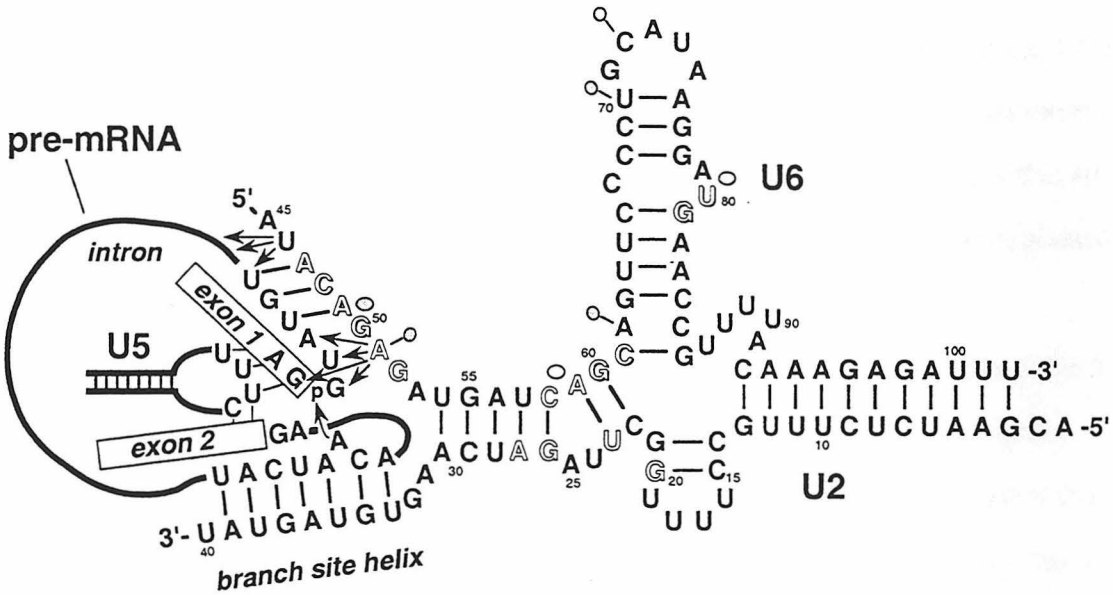
1994) (Brow & Vidaver, 1995)). In addition a sequence near the 3' end of U6 base pairs with the 5' end of U2 ((Datta & Weiner, 1991; Wu & Manley, 1991)) to form a helix important for splicing in mammalian cells, but perhaps not in yeast ((Fabrizio *et al.*, 1989; Madhani *et al.*, 1990). It is not yet clear whether this U2-U6 helix, which does not require the dissociation of U4 to form occurs before or after the U4 destabilization.

The set of interactions between U6 and the 5'-splice site, between U6 and U2 and between U2 and the branch point sequence of the intron serve to hold the pre-mRNA such that the 5' splice site is in close proximity with the branch point sequence (Figure 8). Presumably at this point the catalytic center of the spliceosome is poised to carry out the first step of splicing. Before that can occur an ATP-dependent remodeling of the spliceosome, catalyzed by Prp2 must occur (see below). Following that step the first of the two phosphotransfer reactions takes place.

The set of interactions that occur between U6, U2, U5 snRNAs and the intron not only serve to bring the splice junctions and the branch point sequence into close proximity, the center of this net of interactions also contain the most highly conserved sequences in U2, U5 and U6 and mutations in these sequences block either the first or the second step of splicing. By a complementation assay in which the endogenous U6 RNA in an extract is destroyed and the extract complemented with synthetic U6 RNA it has been possible to explore the role of particular phosphodiester backbone features in splicing. As shown in Figure 8 three different phosphate oxygens in U6 cannot be replaced by sulfur and four different 2' ribose OH groups are essential.

For the second step to occur, requires yet another ATP-dependent remodeling of the spliceosome, catalyzed by Prp16 in conjunction with at least four other proteins. There is considerable evidence that this remodeling changes the nature of the active site. Indeed the two sites must be at least subtly different, because though both are phosphotransfer reactions, the first involves the attack of 2'-OH to form a 2'-5' phosphodiester bond, while the second is an attack of a 3'-OH to yield a 3'-5' phosphodiester bond. Testing of pre-

Figure 8. Summary of the active spliceosome. The lollipops indicate sites where a 2'-deoxy substitutions results in blocks of the first step of splicing. Ovals indicate phosphate oxygens which cannot be replaced by sulfur.


pre-mRNA
intron
exon 1
exon 2
U5
U6
U2
5'
3'-
branch site helix
CAAAGAGAUUU-3'
GUUUCUCUAAGCA-5'

mRNA variants containing nucleoside analogues at the splice sites has provided direct evidence for differences between the chemistry in the two reactions. Incorporation of a 2'-O methyl group at the 5'-splice site had no effect on the first step of splicing while it caused a 15-fold reduction in the rate of the second step (Moore & Sharp, 1992).

In the group I self splicing reaction, the first and second steps are nearly precise forward and reverse reactions. The S_N2 transesterification reaction requires the inversion of the stereochemistry of the reactive oxygen and as expected, only one of the steps, the second, is inhibited by replacing the reactive phosphate with an Rp phosphothiorate diastereomer (Herschlag *et al.*, 1991; Mcswiggen & R.Cech, 1989; Suh & Waring, 1992). By contrast in pre-mRNA splicing, introduction of the Rp phosphothiorate diastereomer at either 5' or 3' splice site blocks splicing at that site while the Sp diastereomer at either site has no effect. Thus the first and second steps of pre-mRNA splicing cannot be explained as forward and reverse reactions catalyzed by the same active site.

Among the interactions which define the second step and the identification of the 3' splice site may be a direct interaction between the 5' G and 3' G residues of the intron (Parker & Siliciano, 1993). Support for this interaction comes from the observation that a GU> AU mutation at the 5' splice site or an AG> AC mutation at the 3' splice site block splicing, the double AU...AC mutant restores splicing to 10% of normal. Thus the putative G-G interaction can be replaced by an A-C interaction. Interestingly, as shall be seen below, a small number of introns actually have the sequence AU.....AC and studies on how this special class of introns are spliced has provided a fascinating and surprising perspective on the importance of the RNA-RNA interactions in the spliceosome.

The AT-AC introns: the exception proves the rule.

A minor class of introns, found so far in only a few Drosophila or metazoan genes has the sequence AT instead of GT at the 5' splice site and AC instead of AG at the 3' splice site. These introns contain idiosyncratic and longer consensus sequences at both 5' and 3' splice sites and in place of the branch point sequence contain the sequence TCCTTAAC 8-11 nucleotides upstream of the 3' splice site. Recently it has been discovered (Hall & Padgett, 1994; Tarn & Steitz, 1996a; Tarn & Steitz, 1996b) that a special spliceosome (atac spliceosme) is assembled to splice these introns. The atac spliceosome has only U5 snRNA in common with the the guag spliceosome (Figure 9). It contains four new snRNPs. U11 takes the place of U1 and recognizes the 5' splice site. U12 takes the place of U2 and pairs with the novel branch point sequence. Two new RNAs, found in low abundance, $U4_{atac}$ and $U6_{atac}$, take the place of U4 and U6. These RNAs exhibit only 40% homology with U4 and U6 RNA. However, $U4_{atac}$ and $U6_{atac}$ pair with each other to form an analogous structure to U4/U6. Most significantly $U6_{atac}$ forms a set of interactions with U12 that are similar to those formed by U6 and U2 in the center of the spliceosome(Figure 10). These include the intermolecular stem Ia and Ib interactions, an intramolecular U6atac stem similar to that seen in U6 following departure of U4. The U6atac-U2 interactions do not include a stem II interaction like that seen between the 5' end of U2 and the 3' end of U6 but recall that this interaction is not essential in yeast. An interaction between U6atac and intron sequences near the 5' splice site brings the 5' splice site close to a polypurine sequence, AAGGAG, the apparent analogue of the highly conserved and essential ACAGAGA sequence in U6. Thus the manner in which these very different RNAs assemble, lends remarkable credence to the importance of the core set of interactions in both spliceosomes.

It seems clear that the two splicing mechanisms are evolutionarily related but why has a distinct mechanism been conserved when it is only being called on to splice a few

Figure 9. The AT-AC snRNPs (from (Tarn & Steitz, 1997)). The Sm-binding sites are indicated by grey boxes. Stretches of four or more nucleotides in U6atac and U4atac that are identical to human U6 or U4 RNA are indicated in purple. Nucleotides predicted to be involved in intermolecular base-pairing interactions with AT-AC introns or other snRNP RNAs are boxed: orange for interactions with the 5' splice site, green for interactions with the branch site, and blue for U6atac-U12 base-pairing.

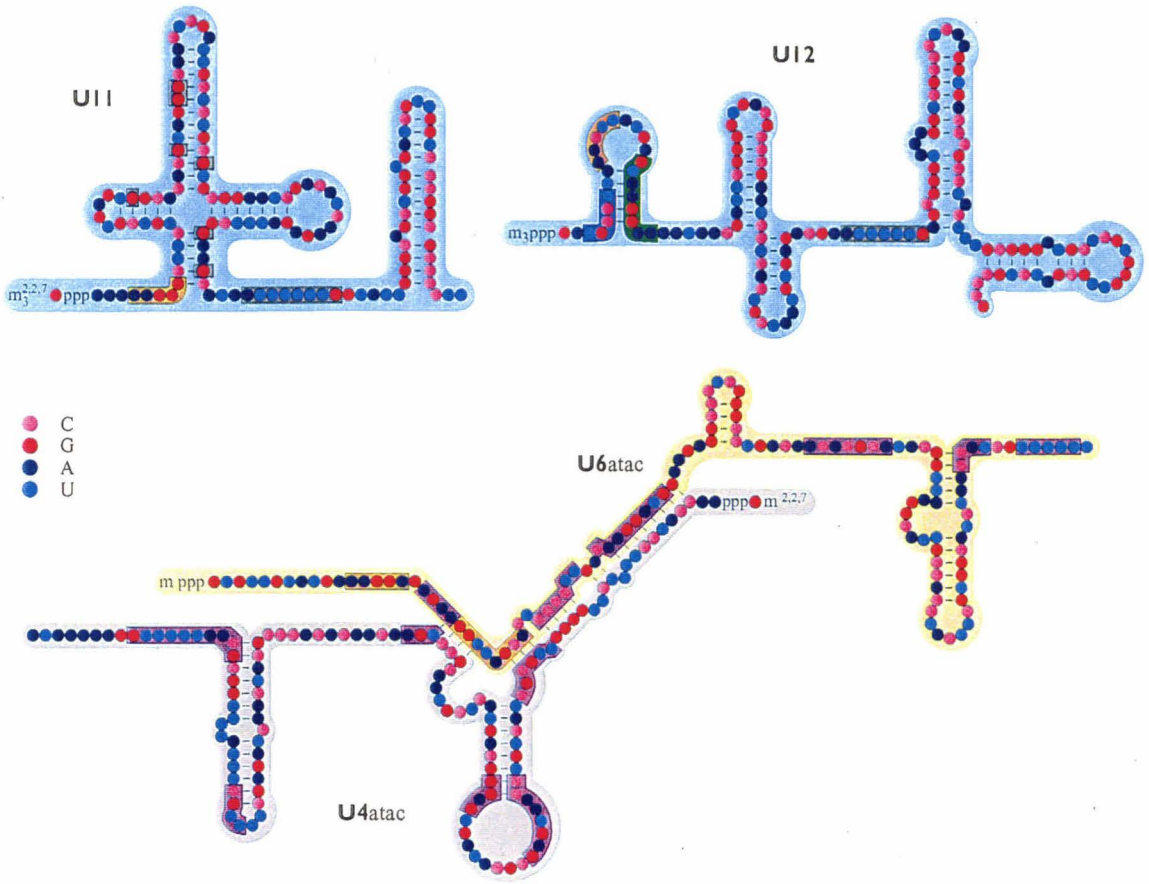
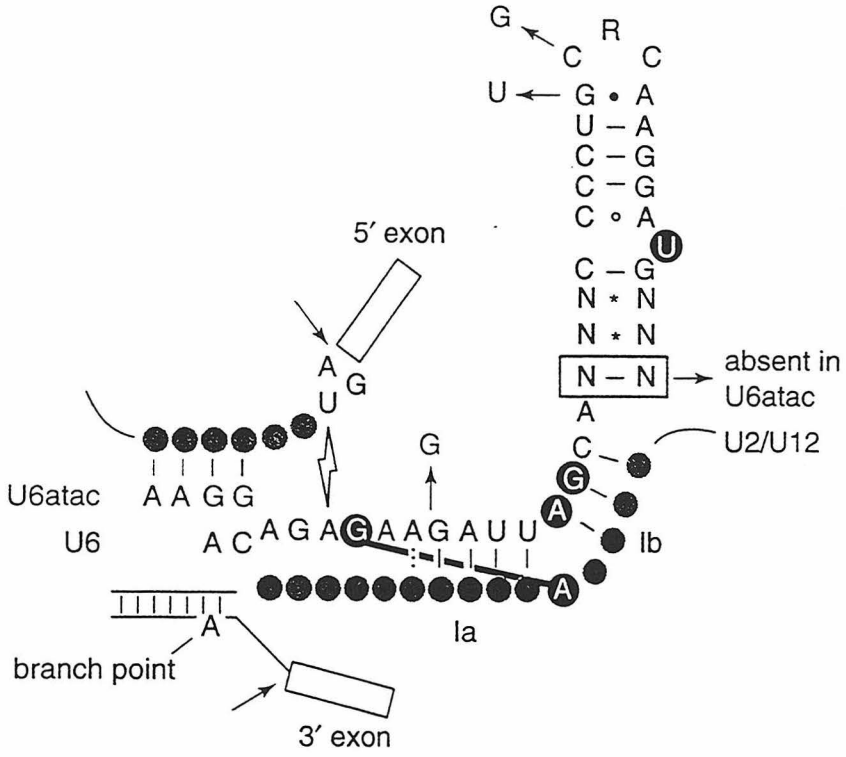

U11
U12
$m_3^{2,2,7}$ ppp
$m_3$ppp
C
G
A
U
U6atac
ppp $m^{2,2,7}$
m ppp
U4atac

Figure 10. AT-AC Spliceosome. RNA-RNA interactions (Tarn & Steitz, 1997). Changes in U6atac RNA relative to U6atac are indicated in green. Red and blue circles indicate bases whose 5' phosphate is important for the first or second step of splicing. A proposed tertiary interaction between the bulge of helix I and the second G of the highly conserved ACAGAG sequence of U6 in yeast is shown in purple. The "lightening bolt' indicates a crosslink established between U6 and the 5' splice site of the major class intron.


G
R
C
C
U
G • A
U — A
C — G
C — G
C ∘ A
U
C — G
N * N
N * N
N — N
absent in
U6atac
A
C
U2/U12
5′ exon
A
G
U
G
U6atac
A A G G
U6
A C
A G A G A A G A U U
G
A
Ib
A
Ia
branch point
3′ exon

introns? As Steven Mount (Mount, 1996) suggested in a commentary on this work: "....ATAC introns are yet another molecular fossil, a way of getting the job done that, like diesel cars or Beta videotapes , is clearly less popular, but whose adherents (people who have diesel cars or Betamax VCRs; genes that contain AT-AC introns) create a continuing demand for the machinery."

The role of proteins in splicing.

We have explored at some length the role of RNA in splicing but it cannot be forgotten that a large number of proteins, perhaps more than 100, are involved in the process and that in many cases they are essential. We know that the proteins play a structural role as snRNP components. They are involved in the early events of splice site recognition and where alternate splice sites are utilized, in splice site selection. Spliceosomal ATPases very likely explain the requirement for hydrolysis of ATP in splicing and finally the close proximity of certain proteins to the center of the spliceosome suggests they could even play a direct or indirect role in catalysis itself. Discussion of two classes of proteins, the SR proteins and the DEAD box spliceosomal ATPases and of one highly conserved protein Prp8 can serve to illustrate what is known and what needs to be discovered about the role of proteins in splicing.

SR proteins

Characterization of the SR proteins began with the discovery that the mammalian protein, ASF/SF2 could, depending on its concentration in extracts, mediate changes between the use of two competing 5' splice sites in early SV40 pre-mRNA (Ge *et al.*, 1991; Krainer *et al.*, 1991). ASF/SF2 contains two N-terminal RNA binding (RRM) domains and a C-terminal domain which contains repeated serine (S) and arginine (R)

residues. The serine residues are frequently phosphorylated and it is likely that phosphorylation controls the activity of the proteins. Subsequently it was discovered that mammalian cells contain a large family of proteins related in structure to ASF/SF2 (Figure 11). Other splicing factors also contain the SR motif including the U1A protein and the 65 and 35kD subunits of U2AF (Zamore *et al.*, 1992; Zhang *et al.*, 1992).

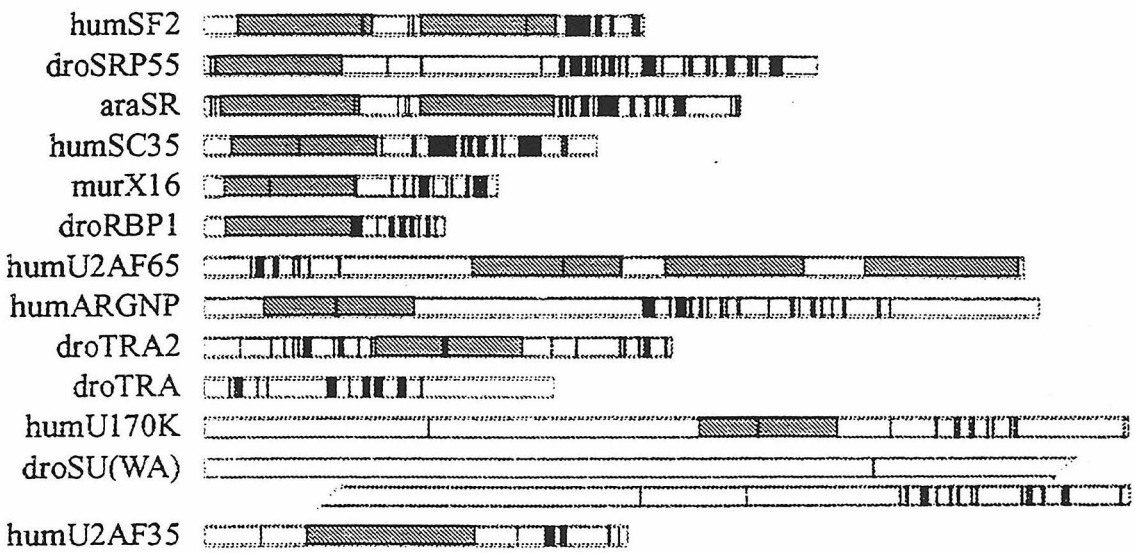
In general SR proteins act early in spliceosome assembly mediating either recognition of the 5' splice site or in the next step of assembly mediating communication between the 5' splice site and the branch point sequence.

SR proteins, for example ASF/SF2 can promote the use of a particular 5' splice site by binding to that site and via interaction with the SR domain of the U170K protein recruit the U1 snRNP to that 5' splice site.

Alternatively the activity of some sub-optimal splice sites can be promoted through binding of one or a set of SR proteins to a purine rich sequence nearby called an exon enhancer. Different exon enhancers may be activated by different sets of SR proteins or by specific SR proteins whose only role is to promote the use of a particular splice site. In Drosophila particular SR proteins, *Tra and Tra-2* control the splicing patterns of the *Double sex* pre-mRNA by binding to an enhancer and recruiting other SR proteins via interaction between the SR domains (Amrein *et al.*, 1994; Boggs *et al.*, 1987; Goralski *et al.*, 1989). This complex alternative splicing pathway ultimately results in sex determination in the fly depending on whether a particular exon is included (female specific *Dsx*) or excluded (male specific *Dsx*).

SR proteins also mediate the early communication between the 5' splice site and the branch point. In mammalian introns the U2AF 65kD protein binds to pyrimidine

Figure 11. SR proteins. Thin vertical lines represent a single RS or SR dipeptide. Grey shading represents RRMs (RNA recognition motifs. (Birney *et al.*, 1993)


humSF2
droSRP55
araSR
humSC35
murX16
droRBP1
humU2AF65
humARGNP
droTRA2
droTRA
humU170K
droSU(WA)
humU2AF35

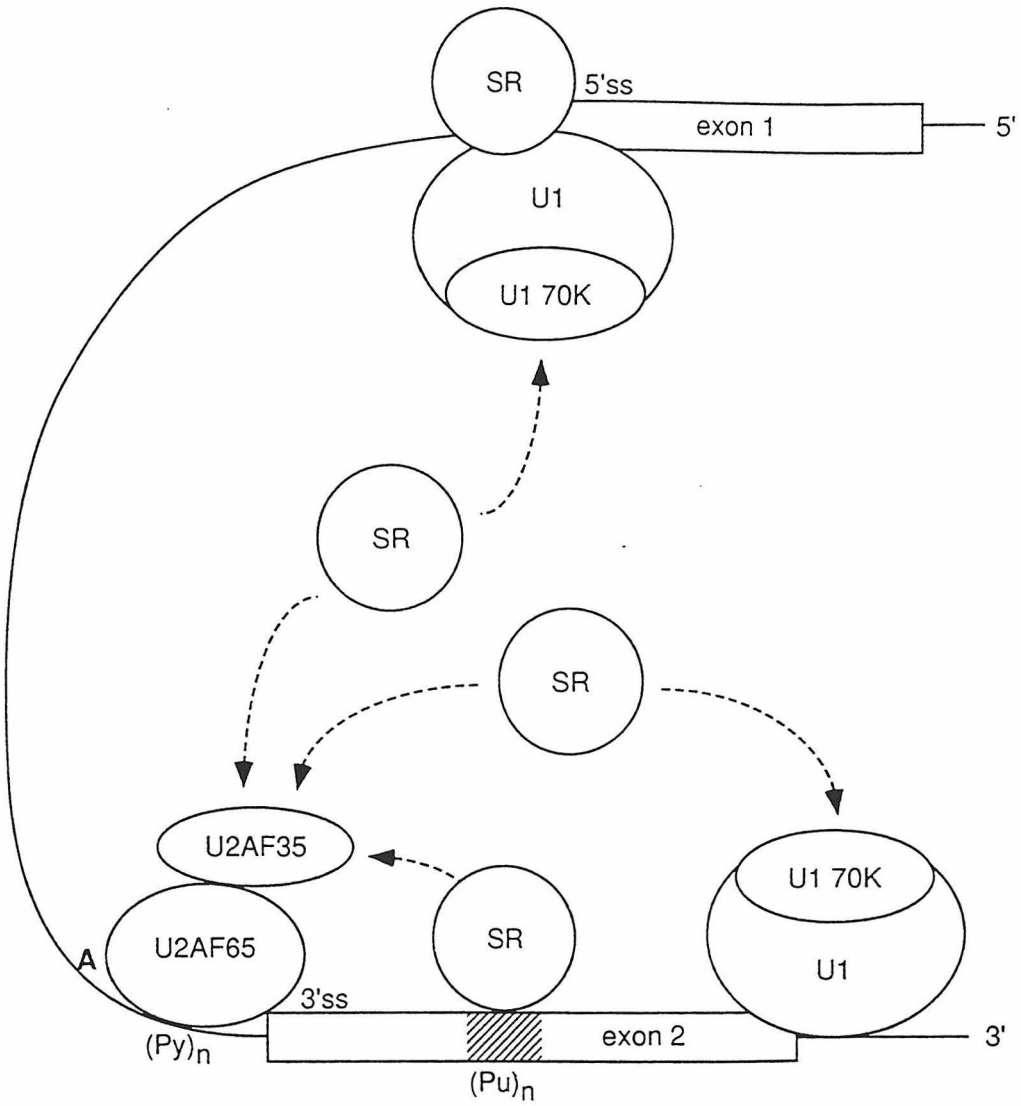
sequences near the branch point (Valcarel et al., 1996). Interestingly, in this case the role of SR domain seems to be promote via arginine-phosphate backbone interactions, the pairing of U2 with the branch point sequence. The 35kD U2AF subunit interacts with U2AF 65kD and via its SR domain with the SR protein SC35. SC35 in turn interacts with the SR domain of U170K (Figure 12) (Fu, 1995). Thus a network of SR interactions serves to provide the crucial early interaction between two adjacent splice sites. This set of interactions probably explains why splicing in mammalian extracts can be made U1 independent through the addition of an excess of SR proteins to the extract. SR proteins, bound near the 5' splice site, could replace the U170K interaction and in subsequent steps U6 could pair with the 5' splice site obviating the U1 displacement reaction which normally occurs at that step (Tarn & Steitz, 1994; Tarn & Steitz, 1995).

The Dead box proteins:

Of particular importance in understanding pre-mRNA splicing is the determination of the role of ATP-hydrolysis in the in vitro reaction. The two phospho-transfer reactions which constitute the steps of splicing would not be expected to require ATP hydrolysis. What then is the requirement for ATP?

Yeast genetic studies of the proteins required for pre-mRNA splicing has revealed at least one source of the ATP requirement in spliceosome assembly: the proteins Prp2, Prp5, Prp16, Prp22 and Prp28, all members of a superfamily whose first member to be identified was the eukaryotic translation initiation factor eIF-4a (Schmid & Linder, 1992). These proteins share a number of sequence motifs including a characteristic ATP binding site with the sequence DEAD/H, hence the family is referred to as the DEAD box proteins. A subclass of DEAD box proteins contain the sequence DEAH and in addition contain an additional 300 amino acid domain at their C termini. DEAD box proteins are ubiquitous in nature and are usually found in multiple copies in all three major lines of descent: the

Figure 12. SR bridging interactions (from (Fu, 1995)). A network of interactions is mediated by SR proteins during early steps of spliceosome assembly. SR proteins may bind directly to the 5' splice site, and thereby mediate pre-mRNA/snRNA interactions defining a functional splice site. SR proteins may also interact directly or indirectly with purine-rich exonic enhancer elements to stimulate and/or stabilize complex assembly at the 3' splice site. Most importantly, SR proteins can simultaneously interact with the U1 70k proteins and the U2AF heterodimer, which may be responsible for bridging the 5' and 3' splice sites and for interactions across exon sequences.


SR
5'ss
exon 1
5'
U1
U1 70K
SR
SR
U2AF35
U2AF65
A
SR
U1 70K
U1
3'ss
exon 2
3'
(Py)n
(Pu)n

bacteria, the archaebacteria and the eukaryotes. There are at least five such proteins in *E. coli*. A recent search of the nearly complete yeast sequence data base found 52 DEAD proteins and 14 DEAH proteins. By extrapolation from the completed 2% of the *C. elegans* genome sequence there are probably at least 100 DEAD proteins in higher eukaryotes. Biochemical studies of several DEAD box proteins have revealed that they are RNA dependent ATPases, and a few have been shown to be RNA helicases because they can separate the strands of double-helical RNA in an ATP dependent reaction. However, such helicase activity has not been demonstrated for most of the DEAD box proteins so far tested, therefore a safer description for the DEAD box proteins required for splicing might be "spliceosomal ATPases."

Studies of the DEAD box proteins suggest that each ATP dependent step in assembly might be mediated by one of these proteins. The formation of the pre-spliceosome requires Prp5 (O'Day *et al.*, 1996; Ruby *et al.*, 1993) . Prp2 is required for an activation step prior to the first step, Prp16 for activation of the second step, and Prp22 for release of the products. Prp2, Prp16 and Prp22 are all members of the DEAH family. What reactions do these proteins catalyze in spliceosome assembly? The assembly pathway is punctuated by a series of RNA rearrangements that are best described as strand displacements or structural isomerizations. A working hypothesis is that each structural isomerization, for example the exchange of U1/5' splice site interactions for U6/5' splice site interactions, is mediated by a specific spliceosomal ATPase. Clearly, to understand exactly what these proteins do and how they work must be a major objective for studies of pre-mRNA splicing.

Prp8

PRP8 is a member of the original set of 10 yeast genes implicated in pre-mRNA. It is a very large protein, 2413 amino acids, and it is highly conserved in sequence between yeast, nematodes and mammals. Originally Prp8 was characterized as an integral U5

snRNP protein (Jackson *et al.*, 1988; Whittaker *et al.*, 1990). The first mutants in PRP8 blocked splicing at the first step. Crosslinks have also been observed between Prp 8 and nucleotides adjacent to the 5' splice site (Reyes *et al.*, 1996; Teigelkamp *et al.*, 1995). It was therefore surprising that a new allele prp8-101, was found to block splicing at the second step (Umen & Guthrie, 1995a). Other alleles of Prp8 affect 3' splice site selection (Umen & Guthrie, 1996) Umen and Guthrie also showed that Prp8 crosslinks to sequences near the 3' splice site (Teigelkamp et al., 1995; Umen & Guthrie, 1995a). Thus this large protein may mediate both steps in splicing and may be close to the active center of the spliceosome. One possibility is that it may serve as a template to insure the correct conformation of the spliceosomal RNA structure, especially the interaction of U5 snRNA with both splice sites (Figure 8). If this is so it may be that the evolution of the splicing mechanism, at first completely independent of proteins as in the Group II introns, is now completely dependent on proteins-a requirement imposed on the system by the great diversity of introns it must splice.

References

Amrein, H., Hedley, M. L. & Maniatis, T. (1994). The role of specific protein-RNA and protein-protein interactions in positive and negative control of pre-mRNA splicing by transformer 2. *Cell* **76**, 735-746.

Ares, M. & Weiser, B. (1995). Rearrangement of snRNA structure during assembly and function of the spliceosome. *Prog. Nucl. Acid Res. and Molec. Biol.* **50**, 131-159.

Beggs, J. D. (1993). Yeast protein splicing factors involved in nuclear pre-mRNA splicing. *Molecular Biology Reports* **18**, 99-103.

Behrens, S. E. & Lührmann, R. (1991). Immunoaffinity purification of a [U4/U6.U5] tri-snRNP from human cells. *Genes Dev.* **5**, 1439-1452.

Berget, S. (1995). Exon recognition in vertebrate splicing. *JBC* **270**(6), 2411-2414.

Birney, E., Kumar, S. & Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *NAR* **21**(25), 5803-5816.

Boggs, R. T., Gregor, P., Idriss, S., Belote, J. M. & McKeown, M. (1987). Regulation of sexual differentiation in *D. melanogaster* via alternative splicing of RNA from *transformer* gene. *Cell* **50**, 739-747.

Brody, E. & Abelson, J. (1985). The "spliceosome": Yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**, 963-967.

Brow, D. A. & Vidaver, R. M. (1995). An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. *RNA* **1**(2), 122-131.

Culver, G. M., McCraith, S. M., Zillmann, M., Kierzek, R., Michaud, N., LaReau, R. D., Turner, D. H. & Phizicky, E. M. (1993). An NAD derivative produced during transfer RNA splicing: ADP-ribose 1"-2" cyclic phosphate. *Science* **261**, 206-208.

Datta, B. & Weiner, A. M. (1991). Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature (Lond.)* **352**, 821-824.

Fabrizio, P., McPheeters, D. S. & Abelson, J. (1989). *In vitro* assembly of yeast U6 snRNP: A functional assay. *Genes Dev.* **3**, 2137-2150.

Fortner, D. M., Troy, R. G. & Brow, D. A. (1994). A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes Dev.* **8**, 221-233.

Frendewey, D. & Keller, W. (1985). Stepwise assembly of a pre-messenger RNA splicing complex requires U-snRNPs and specific intron sequences. *Cell* **42**, 355-367.

Fu, X. D. (1995). The superfamily of arginine serine-rich splicing factors. *RNA* **1**(7), 663-680.

Ge, H., Zuo, P. & Manley, J. L. (1991). Primary structure of the human splicing factor ASF reveals similarities with Drosophila regulators. *Cell* **66**, 373-382.

Goralski, T. J., Edstrom, J. E. & Baker, B. S. (1989). The sex determination locus transformer-2 of Drosophila encodes a polypeptide with similarity to RNA binding proteins. *Cell* **56**, 1011-1018.

Grabowski, P. J., Seiler, S. R. & Sharp, P. A. (1985). A multicomponent complex is involved in the splicing of messenger-RNA precursors. *Cell* **42**, 345-353.

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

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

Hall, S. L. & Padgett, R. A. (1994). Conserved sequences in a class of rare eukaryotic nuclear introns with nonconsensus splice sites. *J Mol Biol* **239**(3), 357-365.

Herschlag, D., Piccirilli, J. A. & Cech, T. R. (1991). Ribozyme-catalyzed and nonenzymatic reactions of phosphate diesters -- rate effects upon substitution of sulfur for a nonbridging phosphoryl oxygen atom. *Biochemistry* **30**(20), 4844-4854.

Jackson, S. P., Lossky, M. & Beggs, J. D. (1988). Cloning of the rna 8 gene of Saccharomyces-cerevisiae, detection of the RNA8-protein, and demonstration that it is essential for nuclear pre-mRNA splicing. *MCB* **8**(3), 1067-1075.

Kim, S. H. & Lin, R. J. (1993). Pre-messenger RNA splicing within an assembled yeast spliceosome requires an RNA-dependent ATPase and ATP hydrolysis. *Proc. Natl. Acad. Sci. USA* **90**, 888-892.

Krainer, A. R., Mayeda, A., Kozak, D. & Binns, G. (1991). Functional expression of cloned human splicing factor SF2: Homology to RNA-binding proteins, U1 70K, and Drosophila splicing regulators. *Cell* **66**, 383-394.

Lesser, C. F. & Guthrie, C. (1993). Mutations in U6 snRNA that alter splice site specificity: Implications for the active site. *Science* **262**, 1982-1988.

Madhani, H. D., Bordonné, R. & Guthrie, C. (1990). Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev.* **4**, 2264-2277.

Madhani, H. D. & Guthrie, C. (1992). A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome. *Cell* **71**, 803-817.

Madhani, H. D. & Guthrie, C. (1994a). Dynamic RNA-RNA interactions in the spliceosome. *Annu. Rev. Genet.* **28**, 677-687.

Mcswiggen, J. A. & R.Cech, T. (1989). Stereochemistry of RNA Cleavage by the *Tetrahymena* ribozyme and evidence that the chemical step is not rate-limiting. *Science* **244**(4905), 679-683.

Moore, M. J., Query, C. C. & Sharp, P. A. (1993). Splicing of precursors to messenger RNAs by the spliceosome. In *RNA World* (Gesteland, R. & Atkins, J., eds.). Cold Spring Harbor Laboratory, New York.

Moore, M. J. & Sharp, P. A. (1992). Site-specifically modified pre-mRNA: Substitutions of the 2'-hydroxyl groups at the splice sites. *Science* **256**, 992-997.

Moore, M. J. & Sharp, P. A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature (Lond.)* **365**, 364-368.

Mount, S. M. (1996). AT-AC introns -- an attack on dogma. *Science* **271**(5256), 1690-1692.

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

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

O'Day, C. L., Dalbadie-McFarland, G. & Abelson, J. (1996). The *Saccharomyces cerevisiae* Prp5 protein has RNA-dependent ATPase activity with specificity for U2 small nuclear RNA. *JBC* **271**(52), 33261-33267.

Parker, R. & Siliciano, P. G. (1993). Evidence for an essential non-Watson-Crick interaction between the 1st and last nucleotides of a nuclear premessenger RNA intron. *Nature* **361**(6413), 660-662.

Peebles, C. L., Gegenheimer, P. & Abelson, J. (1983). Precise excision of intervening sequences from precursor tRNAs by a membrane-associated yeast endonuclease. *Cell* **32**(2), 525-536.

Peebles, C. L., Perlman, P. S., Mecklenburg, K. L., Petrillo, M. L., Tabor, J. H., Jarrell, K. A. & Cheng, H.-L. (1986). A self-splicing RNA excises an intron lariat. *Cell* **44**, 213-223.

Reyes, J., Kois, P., Konforti, B. B. & Konarska, M. M. (1996). The canonical GU dinucleotide at the 5' splice site is recognized by p220 of the U5 snRNP within the spliceosome. *RNA* **2**(3), 213-225.

Ruby, S. W. & Abelson, J. (1991). Pre-mRNA splicing in yeast. *TIGS* **7**, 79-85.

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

Rymond, B. C. & Rosbash, M. (1992). Yeast pre-mRNA splicing. In *Yeast pre-mRNA splicing* (Jones, E. W., Pringle, J. R. & Broach, J. R., eds.). Cold Spring Harbor Laboratory, Cold Spring Harbor.

Schmid, S. R. & Linder, P. (1992). D-E-A-D protein family of putative RNA helicases. *Mol. Microbiol.* **6**, 283-291.

Sontheimer, E. J. & Steitz, J. A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science* **262**, 1989-1996.

Suh, E. & Waring, R. B. (1992). A phosphorothioate at the 3' splice site inhibits the 2nd splicing step in a Group I intron. *NAR* **20**(23), 6303-6309.

Tarn, W. Y. & Steitz, J. A. (1994). SR proteins can compensate for the loss of U1 snRNP functions in vitro. *Genes & Dev* **8**, 2704-2717.

Tarn, W. Y. & Steitz, J. A. (1995). Modulation of 5' splice site choice in pre-messenger RNA by two distinct steps. *PNAS USA* **92**, 2504-2508.

Tarn, W. Y. & Steitz, J. A. (1996a). Highly diverged U4 and U6 small nuclear RNAs required for Splicing Rare AT-AC introns. *Science* **273**(5283), 1824-1832.

Tarn, W. Y. & Steitz, J. A. (1996b). A Novel Spliceosome containing U11, U12, and U5 snRNPs excises a minor class (AT-AC) intron in vivo. *Cell* **84**(5), 801-811.

Tarn, W. Y. & Steitz, J. A. (1997). Pre-mRNA Splicing -- The Discovery of a New Spliceosome Doubles the Challenge. *TIBS* **22**(4), 132-137.

Teigelkamp, S., Newman, A. J. & Beggs, J. D. (1995). extensive interactions of Prp 8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO* **14**(11), 2602-2612.

Umen, J. G. & Guthrie, C. (1995a). Prp16p, Slu7p and Prp8p interact with the 3' splice site in two distinct stages during the second catalytic step of pre-mRNA splicing. *RNA* **1**, 584-597.

Umen, J. G. & Guthrie, C. (1996). Mutagenesis of the yeast gene prp8 reveals domains governing the specificity and the fidelity of 3' splice-site selection. *Genetics* **143**(2), 723-739.

Valcarel, J., Gaur, R. K., Singh, R. & Green, M. R. (1996). Interaction of U2AF(65) RS region with pre-messenger RNA of branch point and promotion of base-pairing with U2 snRNA. *Science* **273**(5282), 1706-1709.

Veen, R. V. d., Arnberg, A. C., Horst, G. V. d., Bonen, L., Tabak, H. F. & Grivell, L. A. (1986). Excised Group II introns in yeast mitochondria are lariats and can be formed by self-splicing in vitro. *Cell* **44**(225-234).

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

Westaway, S. K., Belford, H. G., Apostol, B. L., Abelson, J. & Greer, C. L. (1993). Novel activity of a yeast deletion polypeptide: Evidence for GTP-dependent tRNA splicing. *J. Biol. Chem.* **268**, 2435-2443.

Westaway, S. K., Phizicky, E. M. & Abelson, J. (1988). Structure and function of the yeast tRNA ligase gene. *J. Biol. Chem.* **263**, 3171-3176.

Whittaker, E., Lossky, M. & Beggs, J. D. (1990). Affinity purification of spliceosomes reveals that the precursor RNA processing protein Prp8, a protein in the U5 small nuclear ribonucleoprotein particle, is a component of yeast spliceosomes. *PNAS* **87**, 2216-2219.

Wu, J. & Manley, J. L. (1991). Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature (Lond.)* **352**, 818-821.

Yean, S. L. & Lin, R. J. (1991). U4 small nuclear RNA dissociated from a yeast spliceosome and does not participate in the subsequent splicing reaction. *Mol. Cell. Biol.* **11**, 5571-5577.

Zamore, P. D., Patton, J. G. & Green, M. R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. *Nature (Lond.)* **355**, 609-614.

Zhang, M., Zamore, P. D., Carmo-Fonseca, M., Lamond, A. I. & Green, M. R. (1992). Cloning and intracellular localization øf the U2 small nuclear ribonucleoprotein auxiliary factor small subnuit. *PNAS* **89**, 8769-8773.

Chapter II - Site Specific Deoxynucleotide Substitutions in Yeast U6 snRNA Block Splicing of pre-mRNA *in vitro*

(Chang Hee Kim, Daniel E. Ryan, Tadeusz Marciniec and John Abelson. Published in *EMBO J.* **16**: 2119-2129, 1997)

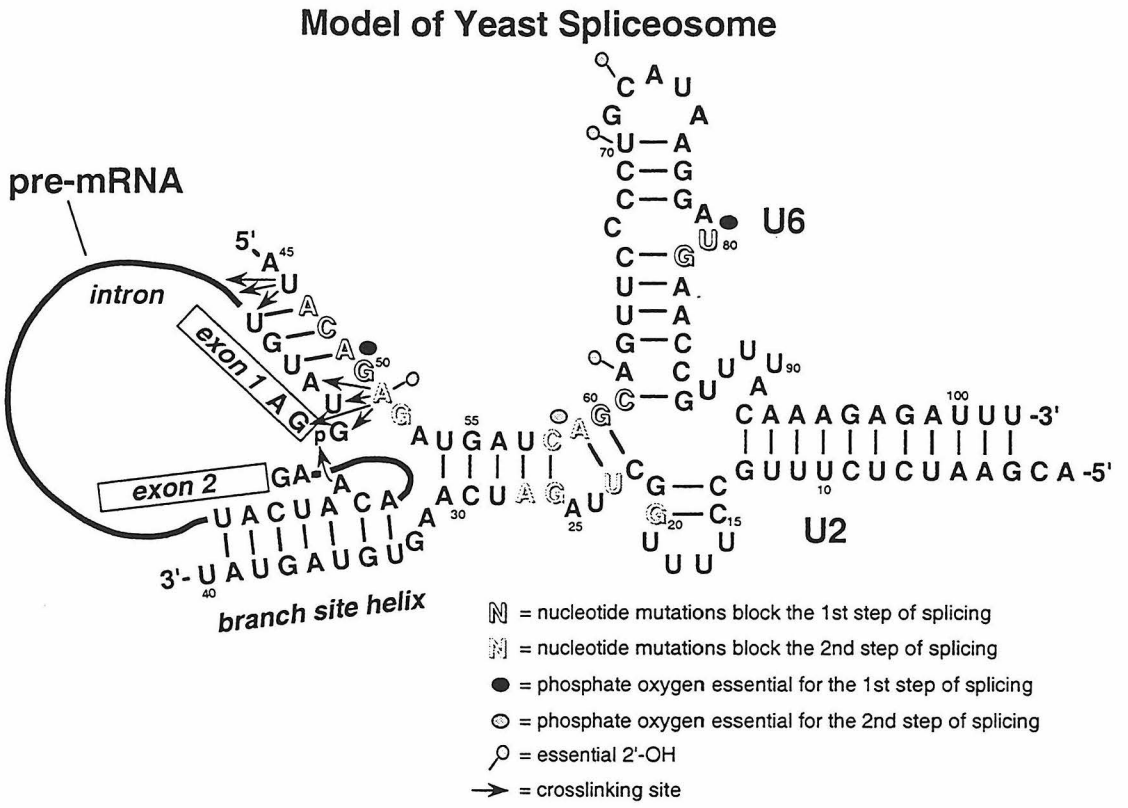
Abstract

We have identified 2'-hydroxyl groups of the U6 phosphate-ribose backbone which are required for reconstitution of splicing activity in U6-depleted yeast extract. To screen the 2'-hydroxyls of yeast U6 at nucleotides 39-88, spanning the conserved central domain, synthetic U6 RNAs were constructed with deoxyribonucleotides incorporated site-specifically. Only four individual deoxynucleotide substitutions blocked splicing activity: dA51 (in the ACAGAG sequence), dA62 (next to the AGC triad), dU70 and dC72 (both in the loop of the 3' intramolecular stem-loop). Native gel analysis revealed that these deoxy-substituted U6 RNAs were competent for assembly of spliceosomes. Interestingly, a 2'-O-methyl substituent at A51, A62, U70, or C72 did not inhibit splicing activity, indicating that the essential 2'-OH groups at these positions in U6 act as hydrogen-bond acceptors or neutral coordinated ligands. The requisite 2'-hydroxyls at A62, U70, and C72 show both similarities and differences relative to the positions of essential 2'-hydroxyls of catalytic domain V of group II ribozymes (Abramovitz et al. 1996). Identification of the essential 2'-hydroxyls at positions 62, 70, and 72 corroborates that the 3' intramolecular stem-loop in U6 plays an important role in pre-mRNA splicing.

Introduction

U6 snRNA is an essential component of fully assembled spliceosomes—large (40S) ribonucleoprotein complexes that mediate the process of pre-messenger RNA splicing (reviews: Ares and Weiser 1995; Madhani & Guthrie 1994b; Sharp 1994; Nilsen 1994). The central domain of U6 (nucleotides 47-85 in yeast) is highly conserved from yeast to mammals (Brow & Guthrie 1988; Roiha et al. 1989; Shumyatsky & Reddy 1993). Two-thirds of the conserved U6 domain is base-paired with U4 snRNA in U4/U6 snRNP complexes (Bringmann et al. 1984; Hashimoto & Steitz 1984; Rinke et al. 1985; Brow & Guthrie 1988; Vankan et al. 1990); the U4/U6 snRNPs deliver U6 to a spliceosome during the ordered process of spliceosome assembly (Pikielny et al. 1986; Cheng & Abelson 1987; Konarska & Sharp 1987). Near completion of assembly, the U4/U6 duplex is dissociated, and U4 may leave the spliceosome (Pikielny et al. 1986; Cheng & Abelson 1987; Lamond et al. 1988; Yean & Lin 1991). Dissociation of U4/U6 allows a 5' region of the U6 conserved domain to base pair with U2 snRNA, thus forming U2/U6 helix I in fully assembled spliceosomes (Madhani & Guthrie 1992) (see model of the fully assembled yeast spliceosome, Figure 1). Concurrently, the 3' region of the U6 conserved domain putatively forms an intramolecular stem-loop as present in noncomplexed U6 snRNPs *in vivo* (Wolff & Bindereif 1991; Fortner et al. 1994; Brow & Vidaver 1995). The conserved ACAGAG region of U6, nts 47-52 in yeast, makes physical contact with the 5' splice site of pre-mRNA, as revealed in crosslinking experiments (Sawa & Shimura 1992; Sawa & Abelson 1992; Wassarman & Steitz 1992; Sontheimer & Steitz 1993; Kim & Abelson 1996). If U6 is simultaneously associated with U2 (at helix I) and pre-mRNA (at the 5' splice site) while U2 is base paired to pre-mRNA at the branch site (Parker et al. 1987), then the pre-mRNA could be folded to juxtapose the 5' splice site and the nucleophilic bulged adenosine of the branch site, perhaps thereby activating the first catalytic step of splicing (cf. Figure 1). Interactions between a conserved stem-loop of

Figure 1. Model of the fully assembled yeast spliceosome. Interactions between U2, U6, and pre-mRNA are illustrated. Mutationally sensitive nucleotides, essential pro-*R* phosphate oxygen atoms, essential 2'-hydroxyl groups (this work), and U6 crosslinking interactions are highlighted.


Model of Yeast Spliceosome
pre-mRNA
intron
exon 1
exon 2
branch site helix
U6
U2
= nucleotide mutations block the 1st step of splicing
= nucleotide mutations block the 2nd step of splicing
= phosphate oxygen essential for the 1st step of splicing
= phosphate oxygen essential for the 2nd step of splicing
= essential 2'-OH
= crosslinking site

U5 snRNA and both the 5' and 3' splice sites are thought to position the 5' and 3' exons for ligation at the second catalytic step of splicing (Newman & Norman 1991, 1992; Wyatt et al. 1992; Sontheimer & Steitz 1993; Frank et al. 1994). The snRNA components of spliceosomes are thought to function as catalysts for the two reactive steps of pre-mRNA splicing. This catalytic RNA hypothesis is founded on the knowledge that the group II self-splicing intron, an RNA-only system (reviewed by Michel & Ferat 1995), has intronic sequences that catalyze self-splicing via the same phosphodiester transfer reactions as occur in pre-mRNA splicing, including the formation of a 2'-5' phosphodiester-linked lariat intermediate.

The structure-function relationships of U6 in the spliceosome are not well understood but are beginning to be elucidated. Specific mutations of individual nucleotides in U6 block the first catalytic step of splicing and prevent spliceosome assembly; other mutations block the second catalytic step (Fabrizio & Abelson 1990, 1992; Vankan et al. 1990; Madhani & Guthrie 1992). In fully assembled spliceosomes, U2/U6 helix I was identified by compensatory mutations and was demonstrated to be essential for splicing activity in yeast (Madhani & Guthrie 1992). Two other U2/U6 helices, flanking the conserved U6 domain, were similarly identified and found to be important for splicing in mammalian cells (Hausner et al. 1990; Wu & Manley 1991; Datta & Weiner 1991; Wolff & Bindereif 1992; Sun & Manley 1995), but are not important in yeast (Fabrizio et al. 1989; Madhani et al. 1990, Yan & Ares 1996). A randomization-selection experiment of both U2 and U6 snRNAs provided evidence of a tertiary contact between U2 and U6 in the helix I region using a method analogous to phylogenetic covariation analysis (Madhani & Guthrie, 1994a). Genetic suppression studies of 5' splice site mutants of pre-mRNA identified base-paired interactions between the 5' splice site and U6 that are important for 5' splice site recognition (Lesser & Guthrie 1993; Kandels-Lewis & Séraphin 1993). U6 structure-function relationships are key features of the fully assembled yeast spliceosome (Madhani & Guthrie 1994b) (cf. Figure 1).

In addition to the U6 sequence and secondary structural requirements for splicing activity, other essential structural features of U6 reside on the phosphate-ribose backbone. Important phosphate oxygen atoms of U6 were identified by incorporating thiophosphate groups at selected positions in the U6 sequence (Fabrizio & Abelson 1992; Yu et al. 1995). For group II ribozymes, the catalytically important phosphate oxygens were similarly identified and suggest that domain V of group II introns has structure-function relationships analogous to the 3' stem-loop of U6 in spliceosomes (Chanfreau & Jacquier 1994). Correlations between pre-mRNA splicing and group II self-splicing are significant because the splicing mechanisms are quite similar in both systems and because the group II system is better understood and could serve as a model for RNA structure-function relationships in spliceosomes. In addition to backbone phosphate groups, specific 2'-hydroxyl groups of ribose rings play critical roles in binding specificity and/or chemical catalysis in self-splicing RNAs and ribozymes (Pyle & Cech 1991; Bevilacqua & Turner 1991; Musier-Forsyth & Schimmel 1992; Pyle et al. 1992; Perreault & Atlman 1992; Herschlag et al. 1993a,b; Smith & Pace 1993; Strobel & Cech 1993; Pley et al. 1994; Abramovitz et al. 1996). These results suggest that specific 2'-hydroxyl groups of U6 may play similar roles.

In the present report, we identified which 2'-hydroxyl groups of U6 are required for the first or second step of splicing in a yeast *in vitro* splicing system (Lin et al. 1985). To screen all 2'-hydroxyls of yeast U6 at nucleotides 39-88, spanning the entire conserved central domain, synthetic U6 RNAs were constructed with deoxynucleotides incorporated site-specifically at single sites or at multiple contiguous sites. Each of the 2'-deoxy U6 RNAs was assayed for reconstitution of splicing activity in U6-depleted yeast extract (Fabrizio et al. 1989), thus the positions of important 2'-hydroxyl groups were revealed by assaying splicing defects for specific 2'-deoxy substituents. This study is the first to investigate deoxynucleotide substitutions in U6 snRNA.

Results

Synthesis of deoxynucleotide-substituted yeast U6 snRNAs

To construct yeast U6 RNAs substituted with a 2'-deoxy- or 2'-O-methyl ribonucleotide at a selected site, each full-length RNA was prepared in pieces, annealed to a complementary single strand of DNA, and ligated using T4 DNA ligase (Kleppe et al. 1970; Moore & Sharp 1992). Oligonucleotides for the central domain were chemically synthesized to incorporate a 2'-deoxy- or 2'-O-methyl ribonucleotide at selected sites from nucleotides 39 through 88. Initially, we prepared U6 RNA in three pieces, nts 1-38, 39-59, and 60-112, such that the 5' and 3' end pieces were transcribed *in vitro* and the central piece (nts 39-59) was chemically synthesized. The yield for this three-piece ligation strategy was as high as 30%. Ligation yields are lower for incorrect 3' and 5' ends at a ligation junction, including the lack of a 5' monophosphate. Indeed, our attempts to transcribe nts 1-58 of U6 (from *Sau*3AI-cut plasmid) primarily produced transcripts that had extra, non-coded 3' terminal nucleotides.

Two improved strategies involved the total chemical synthesis of U6 RNA in either four or five pieces. For the four-piece ligation, synthetic RNA oligonucleotides for yeast U6 nts 1-38, 39-59, 60-79 and 80-112 (all gel purified) were ligated to give full-length U6 in ca. 20% yield. For the five-piece ligation, oligonucleotides for yeast U6 nts 1-38, 39-59, 60-76, 77-94 and 95-112 were ligated in ca. 12% yield. The four-piece ligation was used to incorporate a deoxynucleotide at each individual site from 60 through 79, and the five-piece ligation was used for each site from 80 through 88. We found that large quantities of RNA were best prepared via chemical synthesis of the entire RNA in segments.

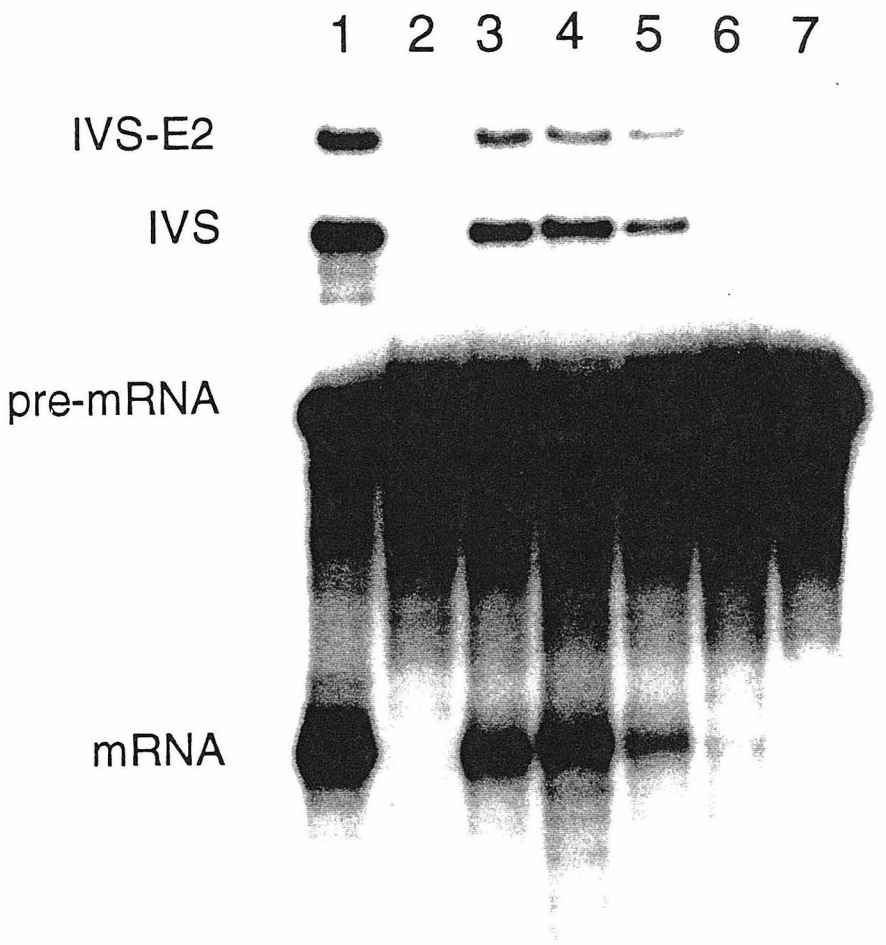
For the all-RNA pieces of U6, gel purification removed nearly all of the synthetic byproducts. To expedite preparation of the 2'-deoxy substituted RNA pieces (29 total), an

improved method for automated RNA synthesis was followed (see Materials & Methods). Synthetic 5'-O-trityl protected oligonucleotides were purified on Nensorb columns to partially remove the short aborts from the chemical synthesis, and the eluted full length 5'-OH oligonucleotides were phosphorylated and used for ligation. For a 0.2 μmole scale automated synthesis, the quantities of Nensorb purified 18-mer and 20-mer products ranged from 1 to 30 nmole. For the five-piece ligations, we started with 40 pmole each of the all-RNA oligonucleotides and 60 pmole of the Nensorb purified, deoxy-substituted oligonucleotide and obtained 2-3 pmole of gel-purified, full-length U6 RNA. In the four-piece ligations, the same amounts of the individual starting materials yielded 6-9 pmole of gel purified full-length U6 RNAs. As less than 50 fmoles of U6 RNA are needed per splicing assay in U6-depleted yeast extract, ample quantities were obtained to conduct the experiments described.

Splicing activity of U6 RNAs substituted with multiple deoxynucleotides in the 5' portion of the central domain

Splicing assays were conducted in yeast extract using [α-^{32}P]uridine-labeled actin pre-mRNA as the substrate. In order to assay U6 RNAs substituted with deoxynucleotides, the endogenous U6 snRNA in the yeast extract was depleted by oligo-directed RNase H digestion using a deoxyoligonucleotide (d1) complementary in sequence to nts 28-54 of yeast U6 RNA (Fabrizio & Abelson 1990). Over the period of incubation, oligonucleotide d1 is itself destroyed by endogenous DNases in the extract. In U6-depleted extract, each of the 2'-deoxy U6 RNAs was assayed for reconstitution of splicing activity (Fabrizio et al. 1989). As a control, wild-type U6 RNA was ligated from three pieces (see above), and the gel-purified, ligated U6 restored splicing activity in U6-depleted extract as efficiently as fully transcribed U6 (lanes 3 and 4, Fig. 2). To rapidly screen the 2'-hydroxyl groups of U6 RNA at nts 39-59, we constructed three U6 RNAs substituted with seven contiguous

Figure 2. Splicing activity of U6-depleted yeast extract reconstituted with U6 RNAs substituted with seven deoxynucleotides at specific contiguous sites. Samples were incubated with [α-^{32}P]uridine-labeled actin pre-mRNA at 23 °C for 30 min, then total RNA was isolated and separated on a denaturing polyacrylamide gel to assay splicing activity: mock-treated extract (lane 1); U6-depleted extract, i.e., extract treated with oligonucleotide d1 to digest endogenous U6 snRNA (lane 2); reconstitution of splicing activity with: *in vitro* transcribed wild-type U6 snRNA (lane 3), wild-type U6 RNA ligated from three pieces (lane 4), ligated U6 RNAs substituted with deoxynucleotides at nts 39-45, 46-52 and 53-59 (lanes 5-7, respectively). Splicing substrate, reaction intermediates and products are labeled as follows: lariat intron-exon 2 (IVS-E2); lariat intron (IVS); pre-mRNA substrate (pre-mRNA); spliced mRNA product (mRNA); exon 1 (E1).


1 2 3 4 5 6 7
IVS-E2
IVS
pre-mRNA
mRNA

deoxynucleotides at specific sites (deoxy box 39-45, 46-52 and 53-59) and assayed each for reconstitution of splicing activity (lanes 5-7, Fig. 2). Only U6 RNA with deoxy box 39-45 (lane 5, Fig. 2) reconstituted splicing activity, although not as efficiently as wild-type U6 RNA (lanes 3 and 4, Fig. 2). Since deoxy box 39-45 was competent for splicing, we did not subsequently make 2'-deoxy substitutions at single sites for nts 39-45. For deoxy boxes 46-52 and 53-59, the inability to reconstitute splicing activity (lanes 6 and 7, Fig. 2) suggested that at least one 2'-OH group in this region is crucial for splicing or that multiple deoxynucleotide substitutions are inhibitory. Subsequently, we prepared U6 RNAs substituted with a single deoxynucleotide at each site from 46 through 59, and the singly substituted U6 RNAs were screened for reconstitution of splicing activity.

Splicing activity of U6 RNAs substituted with a single deoxynucleotide at each site from 46 through 88

Figure 3 shows results of splicing assays for U6 RNAs substituted with a single deoxynucleotide at each site from 46 to 59. For dA51-substituted U6, splicing activity was completely blocked (lane 12, Fig. 3). Interestingly, splicing activity was not inhibited when A51 was replaced by a 2'-O-methyladenosine (lane 13, Fig. 3). The 2'-O-methyl group cannot act as a hydrogen-bond donor, nucleophile, or coordinated anion (in contrast to 2'-OH groups), therefore 2'-O-methyl A51 does not function as such in spliceosomes. Nucleotide A51 is part of the functionally important ACAGAG sequence (Fabrizio & Abelson 1990; Lesser & Guthrie 1993; Kandels-Lewis & Séraphin 1993).

Surprisingly, single deoxynucleotide substitutions at positions 53 to 59 had no effect on splicing (lanes 15-21, Fig. 3), despite the result that U6 with deoxy box 53-59 blocked splicing activity (cf. Figure 2).

Figures 4 and 5 show results of splicing assays for U6 RNAs substituted with a single deoxynucleotide at each site from 60 to 88. In this region, one deoxynucleotide

Figure 3. Splicing activity of U6-depleted yeast extract reconstituted with U6 RNAs substituted site-specifically with a single deoxynucleotide at each site from 46 through 59 (or with a 2'-O-methyl nucleotide at position 51). Samples were incubated with [α-^{32}P]uridine-labeled actin pre-mRNA at 23 °C for 30 min, then total RNA was isolated and separated on a denaturing polyacrylamide gel to assay splicing activity: mock-treated extract (lanes 1 and 8); U6-depleted extract (lanes 2 and 9); reconstitution of splicing activity with: wild-type U6 RNA ligated from three pieces (lanes 3 and 10), ligated U6 RNAs substituted at nts 46-59 with a single deoxynucleotide (lanes 4-7 and 11-21, respectively).

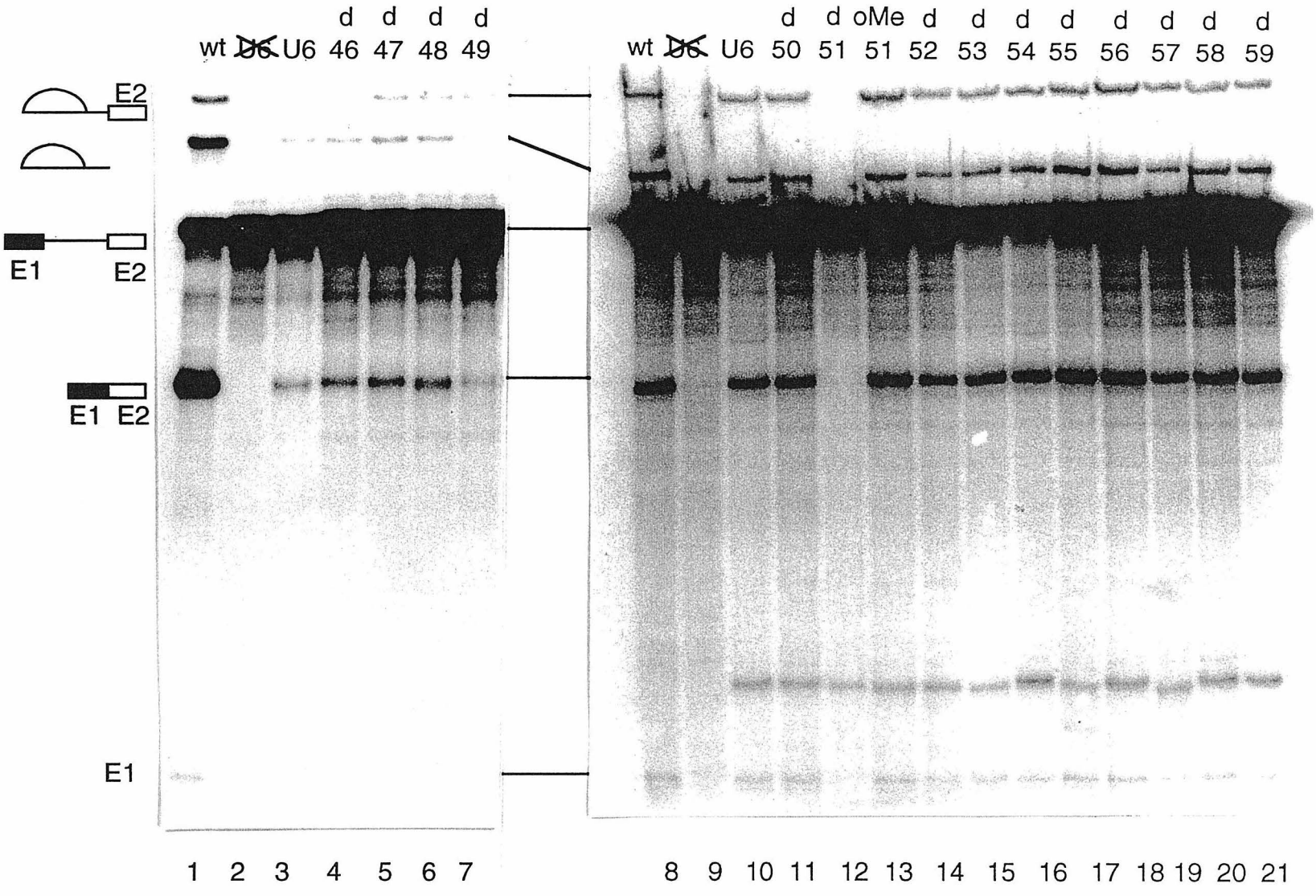

wt
Δ6
U6
d 46
d 47
d 48
d 49
wt
Δ6
U6
d 50
d 51
oMe 51
d 52
d 53
d 54
d 55
d 56
d 57
d 58
d 59
E2
E1
E2
E1 E2
E1
U6
1 2 3 4 5 6 7
8 9 10 11 12 13 14 15 16 17 18 19 20 21

Figure 4. Splicing activity of U6-depleted yeast extract reconstituted with U6 RNAs substituted site-specifically with a single deoxynucleotide at each site from 60 through 73. Samples were incubated with [α-^{32}P]uridine-labeled actin pre-mRNA at 23 °C for 30 min, then total RNA was isolated and separated on a denaturing polyacrylamide gel to assay splicing activity: mock-treated extract (lane 1); U6-depleted extract (lane 2); reconstitution of splicing activity with: wild-type U6 RNA ligated from four pieces (lane 3), ligated U6 RNAs substituted at nts 60-73 with a single deoxynucleotide (lanes 4-17, respectively).

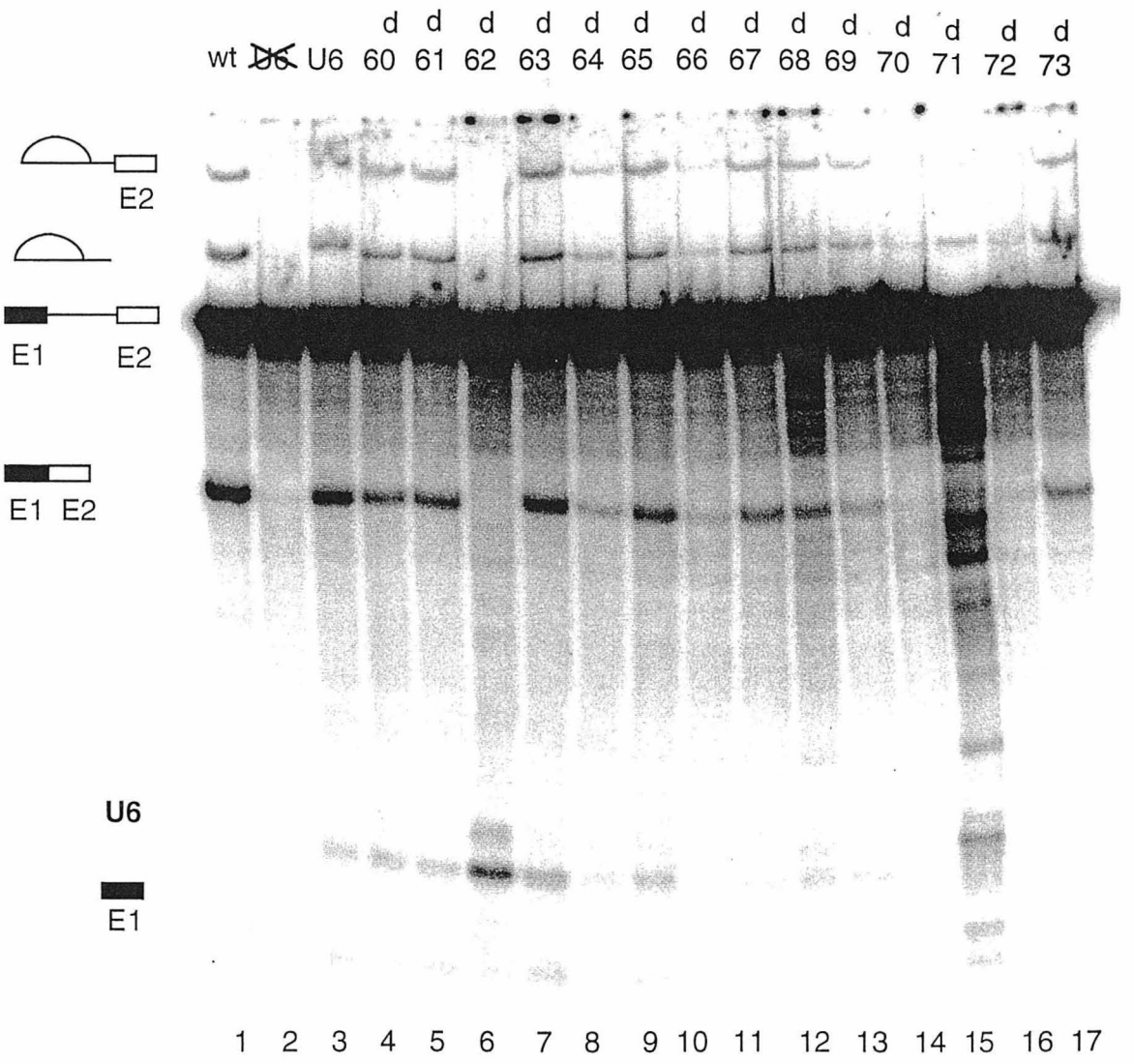
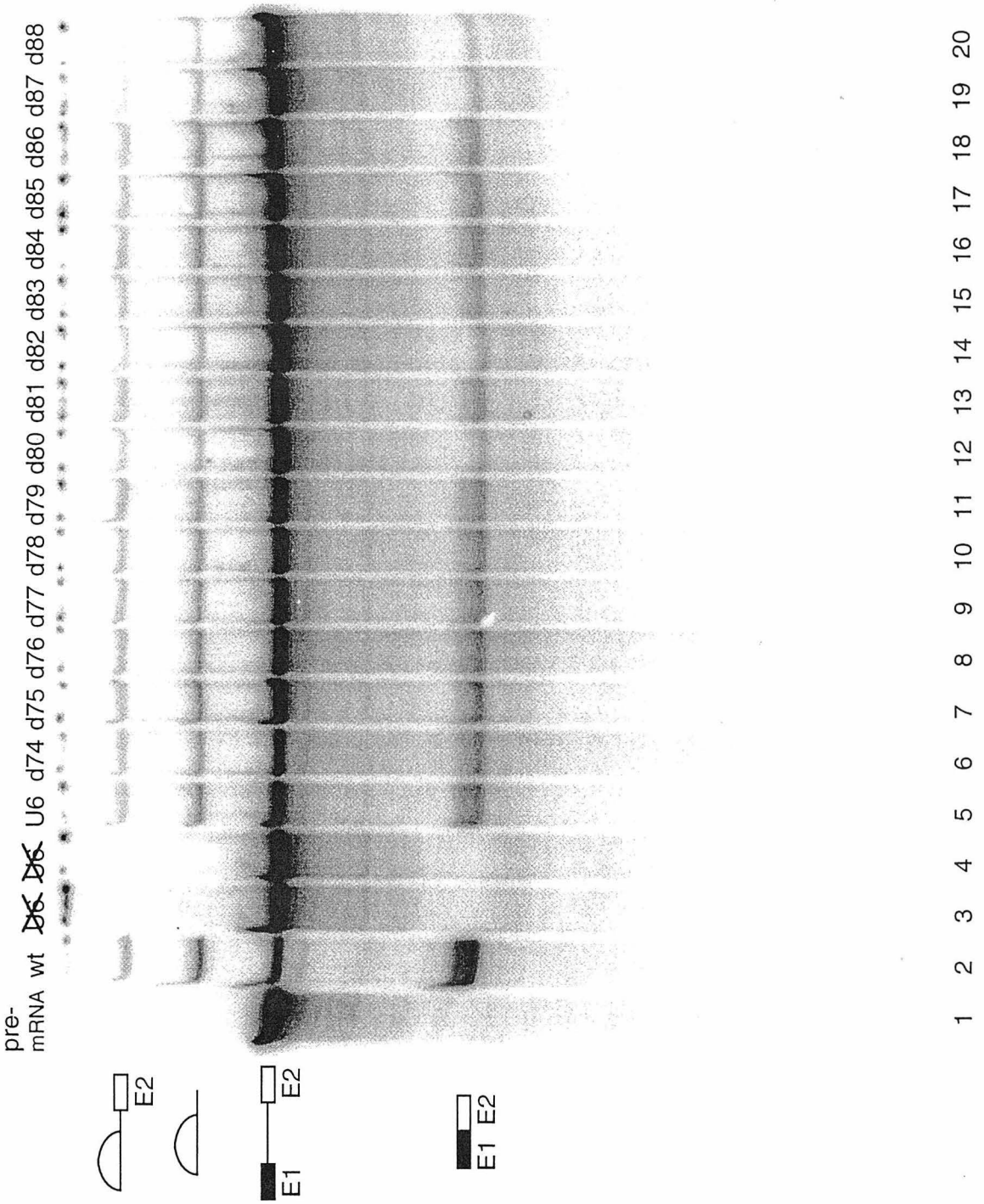

wt
U6
U6
d 60
d 61
d 62
d 63
d 64
d 65
d 66
d 67
d 68
d 69
d 70
d 71
d 72
d 73
E2
E1
E2
E1
E2
U6
E1
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

Figure 5. Splicing activity of U6-depleted yeast extract reconstituted with U6 RNAs substituted site-specifically with a single deoxynucleotide at each site from 74 through 88. Samples were incubated with [α-^{32}P]uridine-labeled actin pre-mRNA at 23 °C for 30 min, then total RNA was isolated and separated on a denaturing polyacrylamide gel to assay splicing activity: actin pre-mRNA substrate (lane 1); mock-treated extract (lane 2); U6-depleted extract (lanes 3 and 4); reconstitution of splicing activity with: *in vitro* transcribed wild-type U6 snRNA (lane 5), ligated U6 RNAs (from five pieces) substituted at nts 74-88 with a single deoxynucleotide (lanes 6-20, respectively).


pre-
mRNA wt U6 d74 d75 d76 d77 d78 d79 d80 d81 d82 d83 d84 d85 d86 d87 d88
E2
E1
E2
E1 E2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

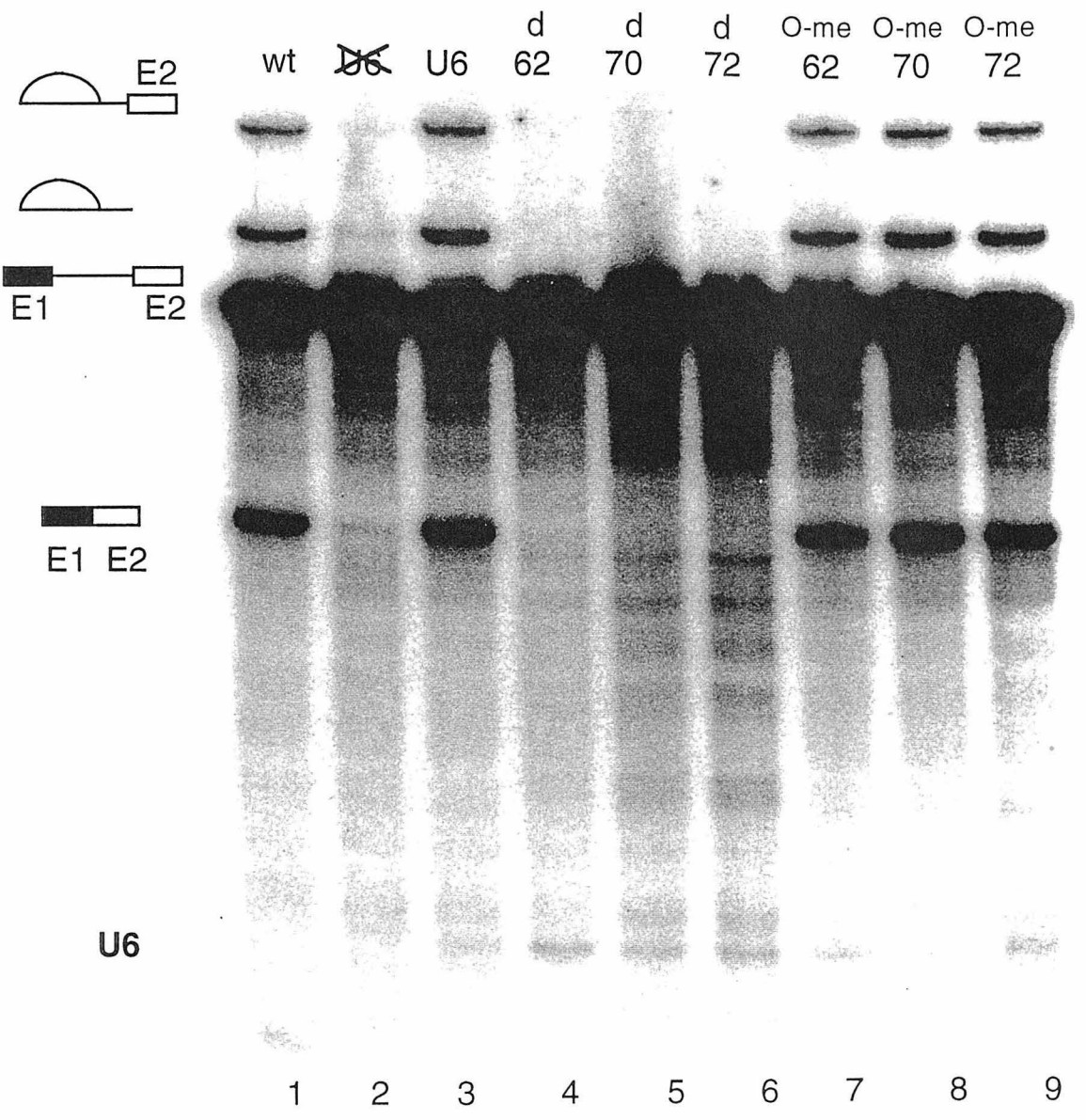
substitution completely blocked splicing, dA62 (lane 6, Fig. 4), and two others showed only traces of spliced products, dU70 (lane 14) and dC72 (lane 16, Fig. 4). Nucleotide A62 is next to the functionally important CAGC sequence, part of U2/U6 helix I (Fabrizio & Abelson 1990; Madhani & Guthrie 1992), and nucleotides U70 and C72 are both part of the loop of the U6 intramolecular 3' stem-loop (Fortner et al. 1994) (cf. Figure 1). As observed for a 2'-O-methyl substituent at nucleotide A51 (see above), a single 2'-O-methyl substituent at A62, U70, or C72 did not inhibit splicing activity (Figure 6), in contrast to the inhibitory effects of deoxynucleotide substitutions at these sites.

Using a set of synthetic oligonucleotide stocks, we were able to reproduce our results. To verify that dA51- and dA62-substituted U6 completely blocked splicing activity and that dU70- and dC72-substituted U6 severely inhibited splicing, we prepared duplicate stocks of the deoxy-substituted oligonucleotides by resynthesizing and purifying them. Using the duplicate stocks, replication of the assay results for dA51-, dA62-, dU70-, and dC72-substituted U6 RNAs (data not shown) confirmed that these individual deoxynucleotide substituents caused the splicing defects observed (cf. Figures 3 & 4).

Native gel analysis of spliceosome assembly in yeast extracts reconstituted with deoxynucleotide-substituted U6 RNAs

For the 2'-deoxy U6 RNAs that blocked splicing activity, it was important to distinguish whether a spliceosome assembly step was blocked or whether spliceosomes were fully assembled and the block occurred at the first catalytic step. Hence, U6 reconstitution experiments which blocked or inhibited splicing were repeated exactly as for the splicing assays except that product mixtures were split in order to assay both splicing activity (on a denaturing gel) and spliceosome assembly (on a native gel). A number of spliceosome assembly intermediates have been separated on native gels and characterized biochemically (Cheng & Abelson 1987; Moore et al. 1993). In yeast, the biochemically

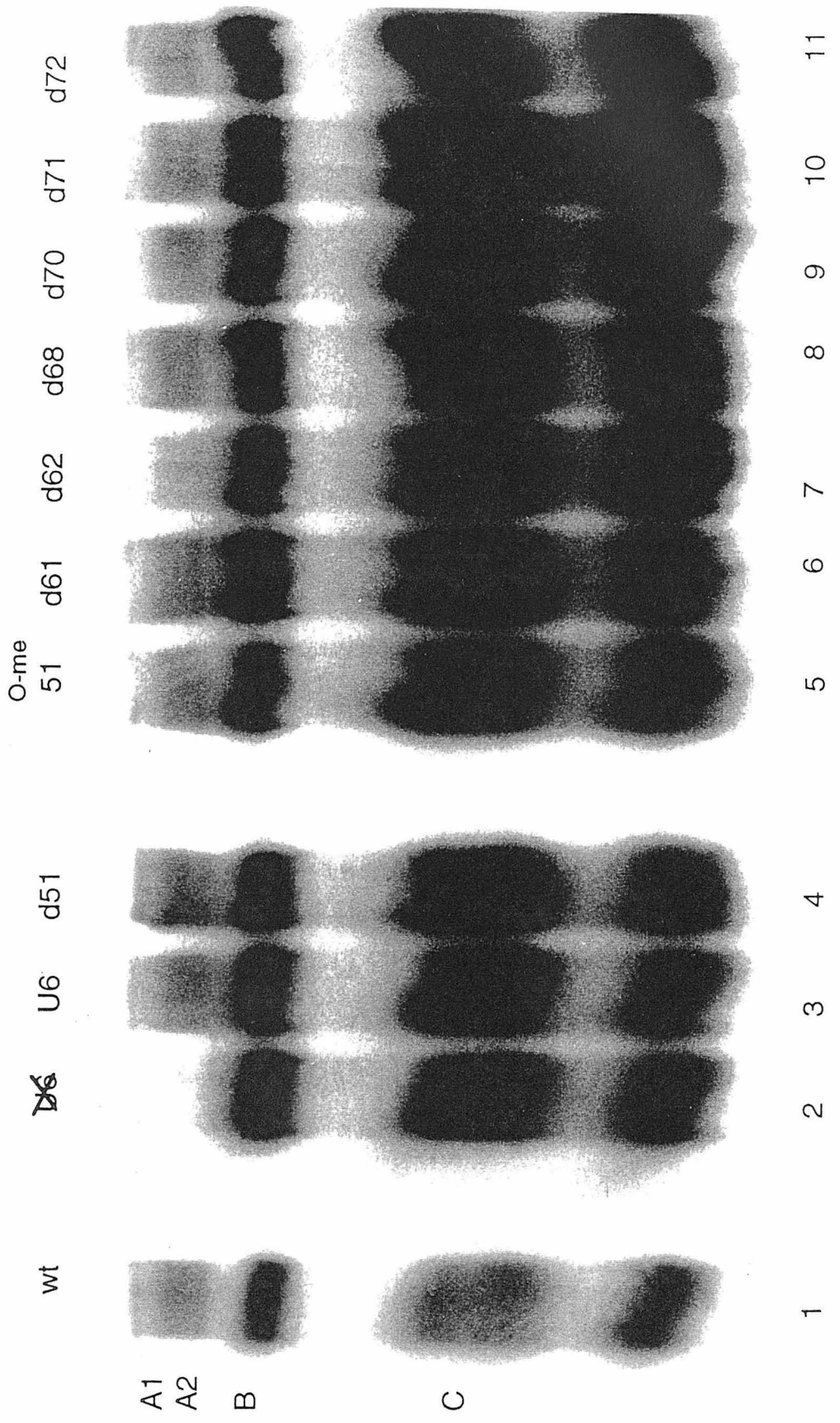
Figure 6. Splicing activity of U6-depleted yeast extract reconstituted with U6 RNAs substituted site-specifically with a single 2'-O-methyl nucleotide at position 62, 70, or 72. Samples were incubated with [α-^{32}P]uridine-labeled actin pre-mRNA at 23 °C for 30 min, then total RNA was isolated and separated on a denaturing polyacrylamide gel to assay splicing activity: mock-treated extract (lane 1); U6-depleted extract (lane 2); reconstitution of splicing activity with: ligated U6 RNAs substituted at positions 62, 70, and 72 with a single deoxynucleotide (lanes 4-6) or a single 2'-O-methyl nucleotide (lanes 7-8).


E2
E1
E2
E1 E2
U6
wt
U6
U6
d 62
d 70
d 72
O-me 62
O-me 70
O-me 72
1 2 3 4 5 6 7 8 9

characterized assembly intermediates are: the "commitment complex" containing pre-mRNA and U1 snRNP, the B complex (U2 snRNP addition), the A2-1 complex (U4/U5/U6 tri-snRNP addition), and the A1 complex (fully assembled spliceosomes). Complexes that contain first-step splicing products (A2-2) and second-step products (A2-3) were identified as well, but these are usually overlapped with A2-1 complexes on native gels. In the standard native gel assay for yeast spliceosome assembly, radiolabeled actin pre-mRNA is the assembly substrate, and partially and fully assembled spliceosomes migrate according to size with the fully assembled A1 complex at the top of the gel. The complexes are identified by inspection of the native gel with reference to a wild-type control lane. Because the A complexes are U6 dependent, they are readily identified via comparison to a U6-depleted control lane. To determine whether the samples generated A2-2 and A2-3 complexes (defined by the presence of spliced intermediates and products), the split portion of each sample was assayed for splicing activity on a denaturing gel.

Figure 7 shows the spliceosome assembly assays for U6 RNAs substituted with a single deoxynucleotide at position 51, 61, 62, 68, 70, 71, or 72 (or with a 2'-O-methyl nucleotide at position 51). Split samples were concurrently assayed for reconstitution of splicing activity (data not shown). The mock-treated yeast extract assembled A and B complexes (lane 1, Fig. 7), and the corresponding splicing assay showed the formation of lariat intermediate and mRNA product indicating that the A2-2 and A2-3 complexes had formed in this sample. When endogenous U6 was depleted by oligo-directed RNase H digestion, no splicing occurred and no A complex was assembled (lane 2, Fig. 7), as expected for complexes that require U6 snRNP for assembly. When wild-type, ligated U6 RNA was added to the same stock of U6-depleted extract as a control, splicing activity and assembly of A complexes were restored (lane 3, Fig. 7). Likewise in the same U6-depleted extract, the deoxynucleotide-substituted U6 RNAs which blocked splicing activity (i.e., dA51-, dA62-, dU70-, and dC72-substituted U6), were all competent for assembly of A complexes (lanes 4-11, Fig. 7). Although assembly of A complexes for dA62-

Figure 7. Spliceosome assembly of U6-depleted yeast extract reconstituted with U6 RNAs substituted site-specifically with a single deoxynucleotide at position 51, 61, 62, 68, 70, 71, or 72 (or with a 2'-O-methyl nucleotide at position 51). Samples were incubated with [α-^{32}P]uridine-labeled actin pre-mRNA at 23 °C for 30 min, treated with heparin, and separated on a native polyacrylamide gel: mock-treated extract (lane 1); U6-depleted extract (lane 2); reconstitution of spliceosome assembly with: wild-type U6 RNA ligated from four pieces (lane 3), ligated U6 RNAs substituted at the positions noted with a single deoxynucleotide (lanes 4-11).


wt
U6
d51
O-me
51
d61
d62
d68
d70
d71
d72
A1
A2
B
C
1
2
3
4
5
6
7
8
9
10
11

substituted U6 was relatively diminished in the experiment shown (lane 7, Fig. 7), replicate experiments (not shown) revealed that dA62 U6 assembled A complexes at relatively high levels. For dA51- and dA62-substituted U6, the absence of spliced intermediates and products on the corresponding denaturing gel (not shown) indicated that the A2-2 and A2-3 complexes were not produced in these samples.

Concentration effects of deoxynucleotide-substituted U6 RNAs in U6-reconstituted yeast extract

We tested various concentrations of U6 RNA in U6-reconstituted samples to determine whether the four inactive deoxy-substituted U6 RNAs (see above) were capable of reconstituting splicing activity at unusually high or low concentrations. Fabrizio et al. (1989) had titrated the concentration of wild-type U6 RNA for reconstitution of splicing activity in U6-depleted extract. As suggested by their titration data, the concentrations of added U6 RNA in our reconstituted samples were in the range of 15-20 nM. In addition to the four inhibitory deoxy-substituted U6 RNAs reported here, we tested the A51U mutant U6 RNA which blocks the second step of splicing and causes an accumulation of A2-2 complexes (Fabrizio & Abelson 1990). As controls for normal assembly and splicing activity, we tested wild-type and dC66-substituted U6 RNAs. Reconstituted samples were split and concurrently assayed for splicing activity and spliceosome assembly (data not shown). For all reconstituted samples, the lowest concentration of added U6 (0.2 nM) produced the lowest levels of reconstituted activity, as expected from the published titration data. As the concentration of added U6 was increased in 10-fold increments (2, 20, and 200 nM), the levels of splicing activity increased for wild-type U6, dC66-substituted U6, and A51U U6 RNAs (this mutant blocked the second step of splicing as expected); however, no splicing activity was observed for dA51-, dA62-, dU70-, and dC72-substituted U6 RNAs over the 10^3-fold range of concentrations. All of the reconstituted

samples showed assembly of complete spliceosomes (A complexes) at increasing levels as the concentration of added U6 was increased, except at the highest level of added U6 (200 nM) which caused a shift from A and B complexes to poorly understood C complexes (Cheng & Abelson 1987). This shift occurred for all U6 RNAs tested, and suggests that the 200 nM level of added U6 was high enough to assess the upper limits of U6 concentration effects. Therefore, the 15-20 nM concentration range for U6 reconstitutions was appropriate, and no unusual concentration effects were found for the four deoxy-substituted U6 RNAs that blocked or severely inhibited splicing activity.

Discussion

In this study of the essential functional groups of the U6 phosphate-ribose backbone, the 2'-OH groups of U6 ribose rings were substituted site-specifically with 2'-H groups to determine which 2'-hydroxyls of U6 are required for splicing in yeast extract. The 2'-hydroxyl group of RNA can contribute binding energy to a bonded interaction or can stabilize a transition state to catalyze a reaction (Pyle & Cech 1991; Bevilacqua & Turner 1991; Musier-Forsyth & Schimmel 1992; Pyle et al. 1992; Perreault & Atlman 1992; Herschlag et al. 1993a,b; Smith & Pace 1993; Strobel & Cech 1993; Pley et al. 1994; Abramovitz et al. 1996). In either case, it can act as a hydrogen-bond donor or acceptor. In binding interactions, the 2'-hydroxyl group can be important for recognition of a ribonucleotide or a structural domain. In catalysis, it can act as an active-site nucleophile, especially when deprotonated (e.g., by a metal hydroxide; Cech 1987), and it can coordinate to active-site metal ions that catalyze a reaction (Pley et al. 1994). Also, 2'-hydroxyls may help organize water molecules for catalysis in an active site. In contrast, a 2'-H group cannot participate in hydrogen bonding or coordinate to metal ions. A deoxyribose ring in DNA conforms to a C2' endo ring pucker, whereas a ribose ring in RNA conforms to a C3' endo ring pucker. The single deoxyribonucleotide substitutions

made in our experiments are not expected to alter the C3' endo conformations of RNA structures; however, the multiple contiguous substitutions may or may not do so (Egli et al. 1993; Ban et al. 1994).

In our initial studies of the 2'-hydroxyls of yeast U6 RNA, we prepared three U6 RNAs substituted with seven contiguous deoxynucleotides (deoxy box 39-45, 46-52, and 53-59). Deoxy box 39-45 reconstituted splicing activity, whereas deoxy boxes 46-52 and 53-59 blocked splicing activity. Surprisingly, single deoxynucleotide substitutions at each site from 53 through 59 did not inhibit splicing. Therefore, deoxy box 53-59 was defective in splicing because it lacked two or more functionally essential 2'-hydroxyl groups. Recently, in the X-ray structure determination of a Group I intron domain (P4-P6), a new binding motif was observed between adjacent phosphate-ribose backbones of the folded RNA which involved pairs of 2'-hydroxyls on contiguous nucleotides (T. R. Cech, personal communication). Such motifs are not necessarily sensitive to a single deoxynucleotide substitution but may be destabilized by multiple substitutions. In our subsequent studies, the 2'-hydroxyl groups of U6 at nucleotides 46-88 were screened individually.

Screens of the fifty 2'-hydroxyls of yeast U6 RNA at nucleotides 39-88, spanning the entire conserved central domain, showed that only two 2'-hydroxyl groups are required for splicing activity: A51 (in the conserved ACAGAG sequence) and A62 (next to the conserved AGC triad; Fig. 1). A deoxynucleotide substitution at either position blocked the first catalytic step of splicing *in vitro*. Two other 2'-hydroxyl groups are required for efficient splicing: U70 and C72 (both in the loop of the 3' intramolecular stem-loop; Fig. 1). A deoxynucleotide substitution at either of these positions severely inhibited splicing, and only traces of spliced products were generated. Native gel analysis revealed that dA51-, dA62-, dU70-, and dC72-substituted U6 RNAs were all competent for complete (or nearly complete) assembly of spliceosomes. The splicing defects for dA51-, dA62-, dU70-, and dC72-substituted U6 RNAs were verified by repeating the chemical syntheses,

ligations, purifications, and splicing assays. Thus, the splicing defects for these four 2'-deoxy U6 RNAs were confirmed. The concentrations of added U6 RNA in U6-reconstituted samples were in the range of 15-20 nM, as suggested by the titration data for U6 reconstitutions (Fabrizio et al. 1989). We tested a 10^3-fold range of concentrations, spanning the titration data of Fabrizio et al., and found no unusual concentration effects for the four inhibitory deoxy-substituted U6 RNAs.

In our experiments to define the chemical roles of the essential 2'-hydroxyl groups of U6, a 2'-O-methyl substituent at nucleotide A51, A62, U70 or C72 did not inhibit splicing activity, indicating that the 2'-OH groups at these positions in U6 do not act as hydrogen-bond donors, nucleophiles or coordinated anions but rather as hydrogen-bond acceptors or a neutral coordinated ligands. (In each case, the presence of the 2'-O-methyl group was confirmed by mass spectrometry.)

In previous studies to identify the functional elements of U6, mutagenesis experiments showed that three sequences in U6 RNA, all in the conserved central domain, are essential for splicing: the ACAGAG sequence (nts 47-52 in yeast), the CAGC sequence (nts 58-61), and the region of the 3' stem-loop bulge (nts 80, 81) (Fabrizio & Abelson 1990, 1992; Vankan et al. 1990; Madhani & Guthrie 1992) (cf. Figure 1). Particular mutations of any one of these nucleotides will inhibit splicing as observed by *in vitro* and *in vivo* splicing assays. Some mutations block splicing activity whereas others specifically block the second step of splicing. In yeast, mutations of A51, G52 (both in the ACAG<u>AG</u> sequence), C58, or C59 (both in the <u>CA</u>GC sequence) are second-step blocks, whereas mutations of the other nucleotides block the first step of splicing. All of the first-step effects were found to block spliceosome assembly at the stage of U4/U5/U6 tri-snRNP addition to the pre-spliceosomal complex of pre-mRNA/U1/U2 (B complex); therefore, the tri-snRNP-containing A complexes were not formed. Mutational analysis of yeast U6 snRNA revealed that nts 54-59 are base paired to U2 snRNA in U2/U6 helix I, an essential interaction for splicing activity *in vivo* (Madhani & Guthrie 1992).

Cooperative with the splicing functions of U6 nucleotide bases are the functions of particular phosphate and ribose groups of the U6 backbone. Backbone phosphate groups of ribozymes and other RNAs provide catalytic metal binding sites as well as binding sites for proteins and other RNA components (Cech 1987). In U6, phosphate oxygen atoms that are required for splicing activity were identified at each of the three essential sequences of U6 (discussed above) by introducing specific phosphorothioate substitutions (Fabrizio & Abelson 1992; Yu et al. 1995). Transcription of phosphorothioate-substituted nucleotides introduces thiophosphate linkages, each with a non-bridging, *(R)*-configured sulfur atom. In yeast U6, a thiophosphate linkage between A49 and G50 or between A79 and U80 blocked splicing activity, just as observed for base mutations of the adjoining nucleotides A49, G50 and U80 (Fig. 1). A thiophosphate linkage between C58 and A59 blocked the second catalytic step of splicing, as observed for base mutations of the adjoining nucleotides C58 and A59 (Fig. 1). The second step was strongly inhibited by a thiophosphate linkage between A51 and G52, nucleotides that specifically block the second step when mutated. The functions of the required phosphate oxygen atoms are strikingly correlated to the functions of the adjoining, essential nucleotide bases.

Comparing our 2'-deoxy substitution results with the results for base mutations and phosphorothioate substitutions in U6 RNA, we note the following: (i) A dA51 substitution in yeast U6 promoted spliceosome assembly but blocked the first step of splicing; therefore, the 2'-hydroxyl group at A51 is essential for splicing catalysis but not for spliceosome assembly. In contrast, base mutations of A51 blocked the second step of splicing. In spliceosomes, nucleotide A51 is held in close proximity to the 5' splice site of pre-mRNA (and to the same site in the lariat intermediate) as demonstrated by crosslinking experiments (Sontheimer & Steitz 1993; Kim & Abelson 1996). (ii) Similar to dA51-substituted U6 RNA, a dA62 substitution in the 3' stem-loop of yeast U6 promoted spliceosome assembly but blocked the first step of splicing, therefore the 2'-hydroxyl group at A62 is essential for splicing catalysis but not for spliceosome assembly.

(iii) Interestingly, the splicing defects observed for dA51- and dA62-substituted U6 RNA appear similar to those observed previously for phosphorothioate-substituted U6 RNAs produced via transcription using α-S GTP or α-S UTP (Fabrizio & Abelson 1992). All of these U6 analogs have modifications in the phosphate-ribose backbone that allow complete (or nearly complete) assembly of spliceosomes but block first-step catalysis. In contrast to the first-step effects of these backbone substituents, all nucleotide base mutations of yeast U6 which blocked the first catalytic step were found to block spliceosome assembly at the stage of U4/U5/U6 tri-snRNP addition (Fabrizio & Abelson 1990). (iv) Yeast U6 with a dU70 or dC72 substitution in the 3' stem-loop promoted spliceosome assembly but severely inhibited the first catalytic step of splicing. Therefore, the 2'-hydroxyls of U70 and C72 are not absolutely required for splicing activity but are essential for splicing efficiency. The 2'-hydroxyl groups of U70 and C72 are the first functional elements identified in the loop region of the U6 3' stem-loop. (v) In contrast to the results for base mutations and phosphorothioate substitutions in U6, no deoxynucleotide substitutions specifically blocked the second catalytic step of splicing. Second-step effects might have been expected for dA51-, dG52-, dC58-, or dA59-substituted U6 (cf. Figure 1). Of these, only dA51-substituted U6 blocked splicing activity and did so just prior to the first catalytic step.

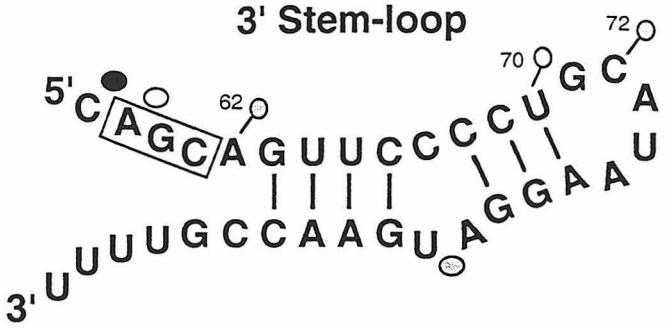
Three of the four 2'-hydroxyl groups identified as functionally essential in yeast U6 are part the 3' stem-loop (Figures 1 & 8). Previously, the only known functional elements of the 3' stem-loop were the nucleotide bases at positions 80 and 81 (at the bulge) and the phosphate group of the adjacent A79-U80 linkage (discussed above). Our results corroborate that the 3' stem-loop plays an important role in splicing catalysis.

Recent results from investigations of phosphorothioate substitutions in U6 RNA and in domain V of group II self-splicing introns suggest that the intramolecular 3' stem-loop of U6 and the domain V stem-loop may be related structurally and/or functionally. The three sites of phosphorothioate inhibition found in yeast U6 (at linkages A49-G50, C58-A59,

Figure 8. A comparison of secondary structure models for the 3' intramolecular stem-loop of yeast U6 (Fortner et al. 1994) and for the domain V stem-loop of group II self-splicing introns (reviewed by Michel & Ferat 1995). The essential 2'-hydroxyl groups identified for yeast U6 (this work) and for the domain V stem-loop (Abramovitz et al. 1996) are illustrated.

Yeast U6

3' Stem-loop



= 2'-OH essential for splicing activity

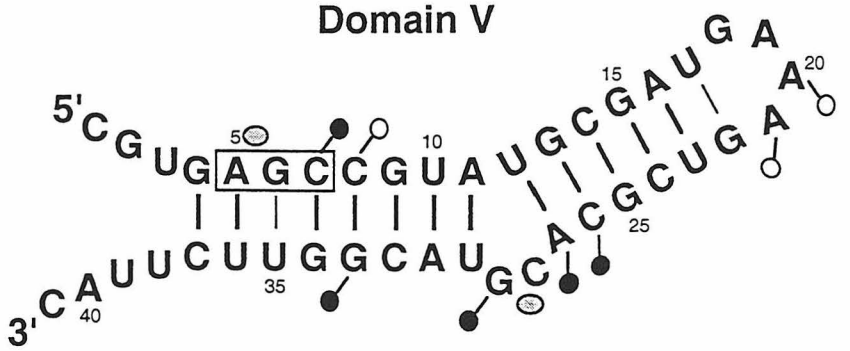
= *(R)*-phosphorothioate blocks splicing activity

= *(R)*-phosphorothioate blocks 2nd step of splicing

= *(R)*-phosphorothioate blocks splicing activity
(identified in ***A. lumbricoides***)

Group II Self-splicing Introns

Domain V



= 2'-OH involved in binding

= 2'-OH involved in chemistry

= *(R)*-phosphorothioate blocks splicing activity

and A79-U80) are equivalent to those found recently in U6 of the nematode *A. lumbricoides* (Fabrizio & Abelson 1992; Yu et al. 1995), and the effects on splicing were the same except that the thiophosphate linkage equivalent to A49-G50 in yeast had a weaker effect in *A. lumbricoides*. In the nematode system, an additional phosphate group was found to participate in splicing—a thiophosphate linkage equivalent to A59-G60 in yeast completely blocked splicing activity. It was known that mutations of the G60 nucleotide in U6 also blocked splicing activity (discussed above). The additional phosphate group identified in nematode U6 may be of general importance in spliceosomes. Unfortunately, the A59-G60 phosphate linkage of yeast U6 could not be screened via phosphorothioate substitution in the yeast experiments (Fabrizio & Abelson 1992). The phosphate groups of nematode U6 equivalent to linkages A59-C60 and A79-U80 in yeast U6 correlate structurally with two phosphate groups found to be essential for the catalytic activity of domain V of group II self-splicing introns (Chanfreau & Jacquier 1994) (Figure 8). These intriguing similarities between the U6 and domain V stem-loops support the hypothesis that both pre-mRNA splicing and group II self-splicing evolved from a common ancestor. Furthermore, the 3' stem-loop of U6 may have structure-function relationships which are analogous to those of the domain V stem-loop (cf. Figure 8).

The noted similarities between the stem-loops of U6 and domain V allow us to compare the essential 2'-hydroxyl groups of the U6 stem-loop with those recently identified in the domain V stem-loop (Abramovitz et al. 1996). In both yeast U6 and domain V stem-loops, the nucleotide linked to the 3' end of the conserved AGC triad, namely A62 in yeast U6 or C8 in domain V, possesses a 2'-hydroxyl group that is essential for splicing activity (Fig. 8). The well-defined kinetic system of the group II ribozyme allowed for the assignment of a binding function to the C8 2'-hydroxyl group. Unfortunately, our splicing assay does not allow us to clearly distinguish between a binding function and a catalytic role. The two other 2'-hydroxyl groups of the U6 stem-loop that severely hindered splicing, at U70 and C72, are located in the loop (Fig. 8). The

2'-hydroxyl groups at these two positions are essential for splicing efficiency and may be important for binding interactions in the spliceosome rather than for catalysis. In domain V, two 2'-hydroxyls in the GNRA-type tetraloop were identified as important for binding (Fig. 8). Although the specific binding interactions of a GNRA tetraloop (Costa & Michel 1995) are expected to differ from those of the pentaloop in the U6 stem-loop, the loop regions in both systems may be important binding domains.

The positions of essential 2'-hydroxyl groups in the yeast U6 stem-loop (i.e., A62 adjoining the AGC triad and loop nts U70 and C72) correlate generally to positions of essential 2'-hydroxyls in domain V; however, several other 2'-OH groups of domain V were found to be important for splicing activity, especially on the 3' side of the stem-loop (Fig. 8). In fact, U6 RNA appears to lack any counterparts to the domain V hydroxyls involved in catalysis. In the domain V studies, deoxynucleotide substitution of a catalytic nucleotide resulted in a 10-20 fold decrease in k_{cat}/K_m, reflecting inhibition of the rate limiting step. Pre-mRNA splicing may have different rate limited steps that are not strongly affected by analogous deoxynucleotide substitutions. The correlations between the functional phosphate groups of U6 and domain V (discussed above) are certainly suggestive that the U6 and domain V stem-loops are comparable structures that may be functionally equivalent. Comparing the essential 2'-hydroxyls of both stem-loops, we note that the few similarities found may point to key features in common for these two systems. Additional information about the tertiary structures and functional groups of U6 and domain V is required before one can make meaningful comparisons. The unknown role of the U6 3' stem-loop is intriguing, and studies are underway to incorporate photocrosslinking nucleotides in this stem-loop to trace its interactions in yeast spliceosomes.

Materials and Methods

Synthesis of U6 RNA fragments

Synthetic RNA oligonucleotides were chemically synthesized on an Applied Biosystems synthesizer model 380A. Nucleotide phosphoramidites were commercially obtained from Glen Research. Oligonucleotides were prepared for wild-type and deoxynucleotide-substituted sequences of U6 RNA: nts 39-59 for the three-piece ligation, nts 1-38, 39-59, 60-79 and 80-112 for the four-piece ligation, and nts 1-38, 39-59, 60-76, 77-94, and 95-112 for the five-piece ligation strategy. For ligation of the RNA pieces of U6 using T4 DNA ligase (see below), complementary DNA oligonucleotides were synthesized to anneal to U6 nts 20-100 (three- and four-piece ligations) and nts 20-112 (five-piece ligation). For the all-RNA pieces of U6, oligonucleotides were purified via denaturing polyacrylamide gel electrophoresis. RNA was visualized by UV shadowing, eluted and ethanol precipitated. For deoxynucleotide substitutions in RNA oligonucleotides 60-79 and 77-94, a new procedure for RNA synthesis (Wincott et al. 1995) was followed using *S*-ethyltetrazole (24 mol-equiv; Aldrich) to activate phosphoramidites and using a new TBDMS-deprotection reagent, a mixture of *N*-methyl-pyrrolidinone (2.25 mL, anhydrous), triethylamine (1 mL) and triethylamine·3 HF (1.5 mL) (Aldrich), combined just prior to use (in the order listed) in a plastic tube kept at 50-60 °C. A 250 µL aliquot of this reagent was added to each tube of base-deprotected oligonucleotide solid. After 1.5 h at 65 °C, the fractions of each oligonucleotide were recombined and diluted with 4 mL of 0.1 M TEAA, pH 7, for purification on Nensorb Prep columns (DuPont) according to the manufacturer's instructions. After lyophilization, the partially purified oligonucleotides were dissolved in 10 mM TrisCl, 1 mM EDTA (or TE) buffer pH 7.5.

Pieces A(1-38) and C(60-112) of U6 RNA for the three-piece ligation were transcribed *in vitro* from a synthetic DNA template with T7 RNA polymerase (cf. Fabrizio

& Abelson 1990). Each transcription mixture (3 mL total) contained 40 mM Tris-HCl pH 8.1, 22 mM $MgCl_2$, 5 mM DTT, 1 mM spermidine, 0.01% triton X-100, 8% polyethylene glycol 8000, 4 mM each NTP, 5 mM DNA template, 660 units/mL RNasin (Promega), 10 units/mL PP_iase and 5000 units/mL T7 RNA polymerase (NEB) and was incubated at 37 °C for 4 h. The C(60-112) transcription mixture was additionally supplemented with 16 mM GMP (to cap the 5' terminus as a 5'-monophosphate for ligation). For full-length U6 RNA, the template for transcription was Bam H1-cleaved pUC18 with the U6 gene (plus an extra coded G at the initiation site) under a T7 promoter. For radiolabeled U6 RNA with low specific activity, a 20 µL reaction mixture contained 1.2 mCi/mL [α-^{32}P]UTP (USB-Amersham) and 3 mM cold UTP. High specific activity [α-^{32}P]uridine-labeled actin pre-mRNA was transcribed *in vitro* using SP6 polymerase (Boehringer Mannheim) as described by Lin et al. (1985) except that the reaction mixture was prepared with 3.4 mCi/mL [α-^{32}P]UTP and 50 µM cold UTP. All transcripts were purified via denaturing polyacrylamide gel electrophoresis. RNA was visualized by UV shadowing, eluted, extracted with phenol/chloroform pH 5.2, and ethanol precipitated.

Analysis of the deoxynucleotide-substituted U6 RNA fragments

After synthesis and partial purification on Nensorb columns, each deoxy-substituted U6 oligonucleotide (nts 39-59, 60-79, or 77-94) was analyzed by mass spectrometry to identify the full-length oligonucleotide and by partial alkaline hydrolysis to verify the position of the deoxynucleotide substituent. Aliquots of the U6 RNA fragments to be analyzed by mass spectrometry were precipitated by adding a 1/4 volume of 3 M ammonium acetate and 3 volumes of ethanol followed by cooling at –20 °C overnight. The RNA pellet was washed with 75% ethanol, redissolved in 200 µL of sterile ddH_2O, and precipitated again as before. The pellet was redissolved in sterile ddH_2O, and 1 µg of each sodium-free oligonucleotide was diluted to 10 µL (ddH_2O) and submitted for analysis

by mass spectrometry. To verify the position of a deoxynucleotide substituent, 40-60 pmole of each deoxy-substituted U6 oligonucleotide (39-59, 60-79, or 77-94) was radiolabeled at the 5' end using [γ-^{32}P]ATP (ca. 15 μCi), T4 polynucleotide kinase (NEB; 1 unit/pmole RNA), 100 μM ATP, 70 mM TrisCl pH 7.5, 10 mM $MgCl_2$, and 5 mM DTT at 37 °C for 1 h. The radiolabeled oligonucleotides were purified via denaturing polyacrylamide gel electrophoresis. After elution from the gel, each sample was treated with 6 μg of carrier tRNA and precipitated. Pellets were dissolved in 0.2 M $NaHCO_3$ buffer pH 9.2 to give ca. 0.25 μM solutions. Aliquots (20 μL) were heated at 90-95 °C for 30 min, then put on ice. A 5 μL portion of each sample was separated on a 20% denaturing polyacrylamide gel to produce a partial hydrolysis ladder with a missing band indicative of the position of the deoxyribonucleotide substituent. Xylene cyanole and bromophenol blue dyes were restricted to lanes without sample.

Ligation of the U6 RNA fragments

Synthetic RNA oligonucleotides require a 5' terminal monophosphate group for ligation. For Figures 3, 4, and 6, each oligonucleotide piece (except the 5' piece) of U6 RNA (100 pmole each piece) was phosphorylated by incubating with 1 unit/pmole T4 polynucleotide kinase (NEB) and 5 μCi [γ-^{32}P]ATP (USB-Amersham) in the supplied buffer (1X: 70 mM Tris-HCl pH 7.6, 10 mM $MgCl_2$, 5 mM DTT) at 37 °C for 30 min, then ATP was added (72 μM) and incubation was continued for an additional 30 min. The specific activity of each phosphorylated oligonucleotide was determined by Cerenkov counting a 1-μL aliquot of the reaction mix (5 μM RNA) spotted on a Whatman DE-81 filter and washed three times in 0.3 M phosphate buffer to remove unincorporated [γ-^{32}P]ATP (Sambrook et al. 1989). For Figures 5, the oligonucleotide pieces (except the 5' piece) of each U6 RNA were phosphorylated as a group (40-60 pmole each piece) using 1

unit/pmole T4 polynucleotide kinase, 0.1 mM ATP, and 0.25 μCi [γ-^{32}P]ATP/pmole RNA in the supplied buffer at 37 °C for 1 h.

For the three- and four-piece ligations, the oligonucleotide pieces of U6 were annealed to a DNA oligonucleotide complementary to U6 nts 20-100 and then ligated by T4 DNA ligase as described by Moore and Sharp (1992), except that RNA oligonucleotides were heated at 90 °C for 2 min for annealing and each ligation mixture (20 μL) contained 66 mM Tris-HCl pH 7.6, 6.6 mM $MgCl_2$, 10 mM DTT, 0.66 mM ATP, 3 μM each RNA piece and cDNA (0.9 mol equiv), and 1 unit/μL T4 DNA ligase (USB-Amersham, high concentration). The product mixtures were extracted with phenol/chloroform pH 5.2 and ethanol precipitated. Full-length ligated U6 RNAs were purified by denaturing polyacrylamide gel electrophoresis. For the five-piece ligations, U6 oligonucleotides in 20 mM Tris-HCl pH 8, 1 mM EDTA solution were annealed to a DNA oligonucleotide complementary to U6 nts 20-112 and then ligated as described by Moore and Sharp, except that the ligation mixture was ca. 0.5 μM for each RNA piece and cDNA (0.9 mol equiv).

U6 reconstitution of splicing activity in yeast extract

Yeast whole-cell extract was prepared according to Lin et al. (1985). To prepare U6-depleted extract, yeast extract was incubated with oligonucleotide d1 (complementary to yeast U6 nts 28-54) to digest endogenous U6 snRNA via endogenous RNase H activity. For U6 depletion, each sample contained 40% yeast extract, 60 mM KPO_4 pH 7, 3 mM $MgCl_2$, 2 mM ATP, 1 mM spermidine, 3% PEG-8000, and 0.3 μM oligonucleotide d1 and was incubated for 30 min at 37 °C, or 40 min at 34 °C. As a control, extract without d1 was heated under parallel conditions. The U6-depleted extract was divided into 5 μL aliquots for the addition of a deoxynucleotide-substituted or wild-type U6 RNA to reconstitute splicing activity as described by Fabrizio et al. (1989). U6 RNA (1 μL of 0.1

μM U6, or ddH_2O for controls) was added to each 5 μL sample, and the reconstituted extracts were incubated at 23 °C for 10 min to assemble U6 snRNPs. Then [α-^{32}P]uridine-labeled actin pre-mRNA (0.5 μL of 2 nM actin in 3× buffer) was added to each sample and incubated at 23 °C for 30 min. The reactions were terminated by digestion and extraction of proteins as described by Lin et al. (1985). The radiolabeled actin pre-mRNA products were separated on a 7 or 8% denaturing polyacrylamide gel (29:1 acrylamide/bisacrylamide) to assay for splicing intermediates and products.

In vitro complex formation

Native polyacrylamide gel electrophoresis was performed as described by Cheng and Abelson (1987) for half of each splicing assay mixture (see above) in order to assay assembly of spliceosomes. Heparin (40 μg per 2 μL of yeast extract) was added to samples reconstituted with deoxy-substituted or wild-type U6 RNA (see above), then samples were loaded directly onto the native gel. Electrophoresis at 200V was conducted at 4 °C for 5 h.

Acknowledgements

We are grateful to Anna Marie Pyle of Columbia University for communication of results prior to publication. We thank Olke Uhlenbeck and Jon Staley for comments on the manuscript. Synthetic oligonucleotides were provided by Nicole Levy, Kelly Redeker-DeWulf, and Suzanna Horvath of the Biopolymer Synthesis group at Caltech. This work was supported by a grant from the National Institutes of Health (GM32637).

References

Abramovitz,D.L., Friedman,R.A. and Pyle,A.M. (1996) Catalytic role of 2'-hydroxyl groups within a group II intron active site. *Science,* **271**, 1410-1413.

Ares,M.,Jr. and Weiser,B. (1995) Rearrangement of snRNA structure during assembly and function of the spliceosome. *Progress in Nucleic Acid Research and Molecular Biology* , **50**, 131-159.

Ban,C., Ramakrishnan,B. and Sundaralingam,M. (1994) A single 2'-hydroxyl group converts *B*-DNA to *A*-DNA. Crystal structure of the DNA-RNA chimeric decamer duplex d(CCGGC)r(G)d(CCGG)with a novel intramolecular G-C base-paired quadruplet. *J. Mol. Biol.*, **236**, 275-285.

Bevilacqua,P.C. and Turner,D.H. (1991) Comparison of binding of mixed ribose-deoxyribose analogues of CUCU to a ribozyme and to GGAGAA by equilibrium dialysis: evidence for ribozyme specific interactions with 2'-OH groups. *Biochemistry* , **30**, 10632-10640.

Bringmann,P., Appel,B., Rinke,J., Reuter,R., Theissen,H. and Lührmann,R. (1984) Evidence for the existence of snRNAs U4 and U6 in a single ribonucleoprotein complex and for their association by intermolecular base pairing. *EMBO J.*, **3**, 1357-1363.

Brow,D. and Guthrie,C. (1988) Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. *Nature*, **334**, 213-218.

Brow,D.A. and Vidaver,R.M. (1995) An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. *RNA* , **1**, 122-131.

Cech,T.R. (1987) The chemistry of self-splicing RNA and RNA enzymes. *Science*, **236**, 1532-1539.

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

Cheng,S.C. and Abelson,J. (1987) Spliceosome assembly in yeast. *Genes & Dev.*, **1**, 1014-1027.

Costa,M. and Michel,F. (1995) Frequent use of the same tertiary motif by self-folding RNAs. *EMBO J.*, **14**, 1276-1285.

Datta,B. and Weiner,A.M. (1991) Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. *Nature*, **352**, 821-824.

Egli,M., Usman, N. and Rich,A. (1993) Conformational influence of the ribose 2'-hydroxyl group: crystal structures of DNA-RNA chimeric duplexes. *Biochemistry*, **32**, 3221-3237.

Fabrizio,P.,McPheeters,D.S. and Abelson,J. (1989) In vitro assembly of yeast U6 snRNP: a functional assay. *Genes & Dev.*, **3**, 2137-2150.

Fabrizio,P. and Abelson,J. (1990) Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. *Science,* **250**, 404-409.

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

Fortner,D.M., Troy,R.G. and Brow,D.A. (1994) A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes & Dev.*, **8**, 221-233.

Frank,D.N., Roiha,H. and Guthrie,C. (1994) Architechture of U5 small nuclear RNA. *Mol. Cell. Biol.*, **14**, 2180-2190.

Hashimoto,C. and Steitz,J.A. (1984) U4 and U6 RNAs coexist in a single small nuclear ribonucleoprotein particle. *Nucleic Acids Res.,* **12**, 3283-3293.

Hausner,T.P., Giglio,L.M. and Weiner,A.M. (1990) Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles. *Genes & Dev.,* **4**, 2146-2156.

Herschlag,D., Eckstein,F. and Cech,T.R. (1993a) Contributions of 2'-hydroxyl groups of an RNA substrate to binding and catalysis by the *Tetrahymena* ribozyme: an energetic picture of an active site composed of RNA. *Biochemistry*, **32**, 8299-8311.

Herschlag,D., Eckstein,F. and Cech,T.R. (1993b) The importance of being ribose at the cleavage site in the *Tetrahymena* ribozyme reaction. *Biochemistry*, **32**, 8312-8321.

Kandels-Lewis,S. and Séraphin,B. (1993) Role of U6 snRNA in 5' splice site recognition. *Science,* **262**, 2035-2039.

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

Kleppe,K., van de Sande,J.H. and Khorana,H.G. (1970) Polynucleotide ligase-catalyzed joining of deoxyribo-oligonucleotides on ribonucleotide templates and ribo-oligonucleotides on deoxyribonucleotide templates. *Proc. Natl. Acad. Sci. USA* , **67**, 68-73.

Konarska,M.M. and Sharp,P.A. (1987) Interactions between small ribonucleoprotein particles in formation of spliceosomes. *Cell*, **49**, 763-774.

Lamond,A.I., Konarska,M.M., Grabowski,P.J. and Sharp,P.A. (1988) Spliceosome assembly involves the binding and release of U4 small nuclear ribonucleoprotein. *Proc. Natl. Acad. Sci. USA* , **85**, 411-415.

Lesser,C.F. and Guthrie,C. (1993) A novel base-pairing interaction between U2 and U6 snRNAs suggest a mechanism for catalytic activation of the spliceosome. *Science,* **262**, 1982-1988.

Lin,R.J., Newman,A.J., Cheng,S.-C. and Abelson,J. (1985) Yeast pre-mRNA splicing in vitro. *J. Biol. Chem.*, **260**, 14780-14792.

Madhani,H.D., Bordonné,R. and Guthrie,C. (1990) Multiple roles for U6 snRNA in the splicing pathway. *Genes & Dev.* **4:** 2264-2277.

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.

Madhani,H.D. and Guthrie,C. (1994a) Randomization-selection analysis of snRNAs in vivo: evidence for a tertiary interaction in the spliceosome. *Genes & Dev.*, **8**, 1071-1086.

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

Michel,F. and Ferat,J.-L. (1995) Structure and activities of group II introns. *Annu. Rev. Biochem.*, **64**, 435-461.

Moore,M.J. and Sharp,P.A. (1992) Site-specifically modified pre-mRNA: substitutions of the 2'-hydroxyl groups at the splice sites. *Science*, **256**, 992-997.

Moore,M.J., Query,C.C., and Sharp,P.A. (1993) Splicing of precursors to messenger RNAs by the spliceosome. In Gesteland,R. and Atkins,J. (eds.), *The RNA World*, New York: Cold Spring Harbor.

Musier-Forsyth,K. and Schimmel,P. (1992) Functional contacts of a transfer RNA synthetase with 2'-hydroxyl groups in the RNA minor groove. *Nature*, **357**, 513-515.

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

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

Parker,R., Siliciano,P.G. and Guthrie,C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell*, **49**, 229-239.

Perreault,J.-P. and Atlman,S. (1992) Important 2'-hydroxyl groups in model substrates for M1 RNA, the catalytic RNA subunit of RNase P from *E. Coli. J. Mol. Biol.*, **226**, 399-409.

Pikielny,C.W., Rymond,B.C. and Rosbash,M. (1986) Electrophoresis of ribonucleoproteins reveals an ordered assembly of yeast splicing complexes. *Nature*, **324**, 341-345.

Pley,H.W., Flaherty,K.M. and McKay,D.B. (1994) Three-dimensional structure of a hammerhead ribozyme. *Nature*, **372**, 68-74.

Pyle,A.M. and Cech,T.R. (1991) Ribozyme recognition of RNA by tertiary interactions with specific ribose 2'-OH groups. *Nature,* **350**, 628-631.

Pyle,A.M., Murphy,F.L., and Cech,T.R. (1992) Ribozyme substrate binding site in the catalytic core of the *Tetrahymena* ribozyme. *Nature,* **358**, 123-128.

Rinke,J., Bernd,A., Digweed,M. and Lührmann,R. (1985) Localization of a base-paired interaction between small nuclear RNAs U4 and U6 in intact U4/U6 ribonucleoprotein particles by psoralen cross-linking. *J. Mol. Biol.,* **185**, 721-731.

Roiha,H., Shuster,E.O., Brow,D.A. and Guthrie,C. (1989) Small nuclear RNAs from budding yeasts: phylogenetic comparisons reveal extensive size variation. *Gene,* **82**, 137-144.

Sambrook,J., Fritsch,E.F., Maniatis,T. (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory Press, New York, USA.

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

Sawa,H. and Shimura,Y. (1992) Association of U6 snRNA with the 5'-splice site region of pre-mRNA in the spliceosome. *Genes & Dev.*, **6**, 244-254.

Sharp,P. (1994) Split genes and RNA splicing (Nobel lecture). *Angew. Chem.Int. Ed. Engl.,* **33**, 1229-1240.

Shumyatsky,G. and Reddy,R. (1993) Compilation of small RNA sequences. *Nucleic Acids Res.*, **21**, 3017.

Smith,D. and Pace,N.R. (1993) Multiple magnesium ions in the ribonuclease P reaction mechanism. *Biochemistry,* **32**, 5273-5281.

Sontheimer,E.J. and Steitz,J.A. (1993) The U5 and U6 small nuclear RNAs as active site components of the spliceosome. *Science*, **262**, 1989-1996.

Strobel,S.A. and Cech,T.R. (1993) Tertiary interactions with the internal guide sequence mediate docking of the P1 helix into the catalytic core of the *Tetrahymena* ribozyme. *Biochemistry*, **32**, 13593-13604.

Sun,J.-S., and Manley,J.L. (1995) A novel U2-U6 snRNA structure is necessary for mammalian mRNA splicing. *Genes & Dev.*, **9**, 843-854.

Vankan,P., McGuigan,C. and Mattaj,I.W. (1990) Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation in *Xenopus* oocytes. *EMBO J.*, **9**, 3397-3404.

Wassarman,D.A. and Steitz,J.A. (1992) Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. *Science,* **257**, 1918-1925.

Wincott,F., DiRenzo,A., Shaffer,C., Grimm,S., Tracz,D., Workman,C., Sweedler,D., Gonzalez,C., Scaringe,S. and Usman,N. (1995) Synthesis, deprotection, analysis and purification of RNA and ribozymes. *Nucl. Acids Res.*, **23**, 2677-2684.

Wolff,T. and Bindereif,A. (1991) Conformational changes of U6 RNA during the spliceosome cycle: an intramolecular helix is essential both for initiating the U4-U6 interaction and for the first step of splicing. *Genes & Dev.,* **7**, 1377-1389.

Wolff,T. and Bindereif,A. (1992) Reconstituted mammalian U4/U6 snRNP complements splicing: a mutational analysis. *EMBO J.*, **11**, 345-359.

Wu,J.A. and Manley,J.L. (1991) Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature*, **352**, 818-821.

Wyatt,J.R., Sontheimer,E.J. and Steitz,J.A. (1992) Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes & Dev.*, **6**, 2542-2553.

Yan,D. and Ares,M.,Jr. (1996) Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits. *Mol. Cell. Biol.*, **16**, 818-828.

Yean,S.L. and Lin,R.J. (1991) U4 small nuclear RNA dissociates from a yeast spliceosome and does not participate in the subsequent splicing reaction. *Mol. Cell. Biol.*, **11**, 5571-5577.

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

Chapter III: Site-specific crosslinks of yeast U6 snRNA to the pre-mRNA near the 5'-splice site.

(Chang Hee Kim and John Abelson. Published in *RNA* **2**, 995-1010)

Abstract

We have introduced a single photochemical crosslinking reagent into specific sites in the central domain of U6 to identify the sites that are in close proximity to the pre-mRNA substrate. Four distinct U6 snRNAs were synthesized with a single 4-thiouridine (4-thioU) at positions 46, 51, 54 and 57, respectively. Synthetic U6 RNA containing the 4-thioU modifications can functionally reconstitute splicing activity in cell free yeast splicing extracts depleted of endogenous U6 snRNA. Upon photoactivation with UV (> 300nm), 4-thioU at position 46 forms crosslinks to pre-mRNA near the 5' splice site at nucleotides +4, +5, +6, and +7 in the intron, whereas 4-thioU at position 51 crosslinks to the pre-mRNA at positions -2, -1, +1, +2, +3, and at the invariant G in the lariat intermediate. All crosslinks are dependent on the presence of ATP and the splicing substrate. The two crosslinks to the pre-mRNA from position 46 and 51 of U6 can also occur in prp2 heat-inactivated yeast splicing extracts blocked immediately prior to the first chemical step. Significantly, the crosslink from position 51 can undergo subsequent splicing when the mutant extract is complemented with functional Prp 2 protein in a chase experiment, indicating that the crosslink reflects a functional interaction that is maintained during the first step. The crosslink to lariat intermediate appears when the mutant spliceosomes are complemented with functional Prp2 protein added exogenously. This experiment is a paradigm for future studies in which different mutant extracts are used to establish the stage in assembly at which particular RNA-RNA interactions defined by unique crosslinks occur.

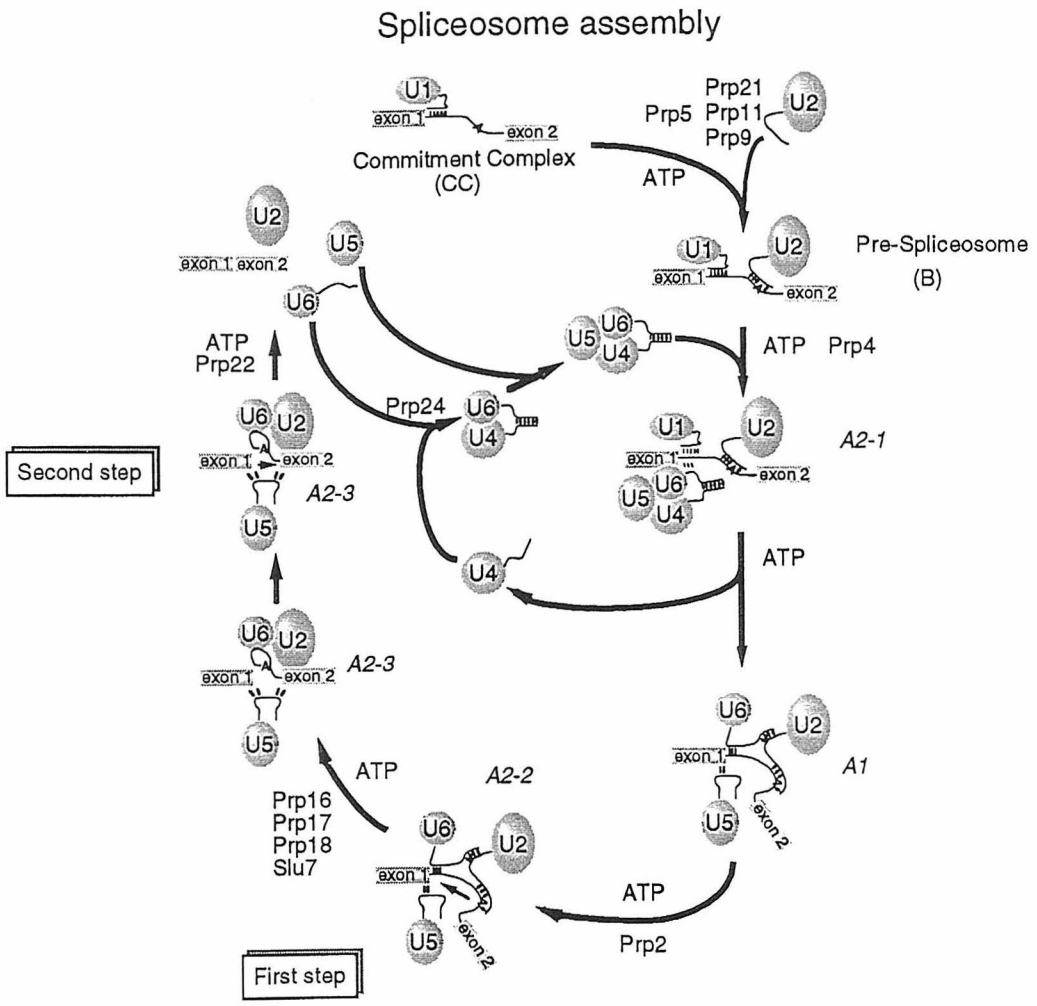
Introduction

The intervening sequences (introns) that interrupt eukaryotic genes are excised and the coding sequences (exons) spliced post-transcriptionally, at the RNA level. The precursor messenger RNA (pre-mRNA) assembles with numerous proteins and five small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) to form a dynamic, enzymatic complex (the spliceosome) which catalyzes the splicing reaction. Two consecutive phosphodiester bond transesterifications with inversion of stereochemical configuration are required for the net chemical steps in exon splicing. First, a 2'-OH of a conserved intronic adenosine carries out an SN2 nucleophilic attack on the phosphate at the 5' exon-intron junction (5' splice site)-- creating a looped structure (lariat intermediate) closed by a novel 2'-5' linkage and releasing the 5'exon as a leaving group. Second, the 3'-OH of the free 5'-exon acts as the nucleophile which forms a bond to the phosphate at the intron-3' exon junction (3' splice site) resulting in ligation of the two exons and release of the intron as a lariat product. Surprisingly, the splicing reaction in Group II RNA introns occurs by the same net chemical mechanism but does not require exogenous protein or ATP in vitro. Thus, the prevailing hypothesis is that the RNA components of the spliceosome constitute its chemically active structure.

This hypothesis is strongly supported by biochemically and genetically well-characterized interactions between conserved nucleotides of the splicing substrate and the snRNAs (for recent reviews, refer to Ares and Weiser, 1995; Madhani and Guthrie, 1994; Newman, 1994; Nilsen, 1994; Legrain and Chanfreau, 1994). These interactions occur during the ordered assembly of the spliceosome and involve a series of structural isomerizations in which initial base-pairing interactions are replaced by new interactions.

Spliceosome assembly, diagrammed in Figure 1A, begins with the interaction of the U1 snRNP with the pre-mRNA. The sequence at the 5' end of U1 RNA is complementary to conserved intron nucleotides at the 5' splice site, and base-pairing of these regions is

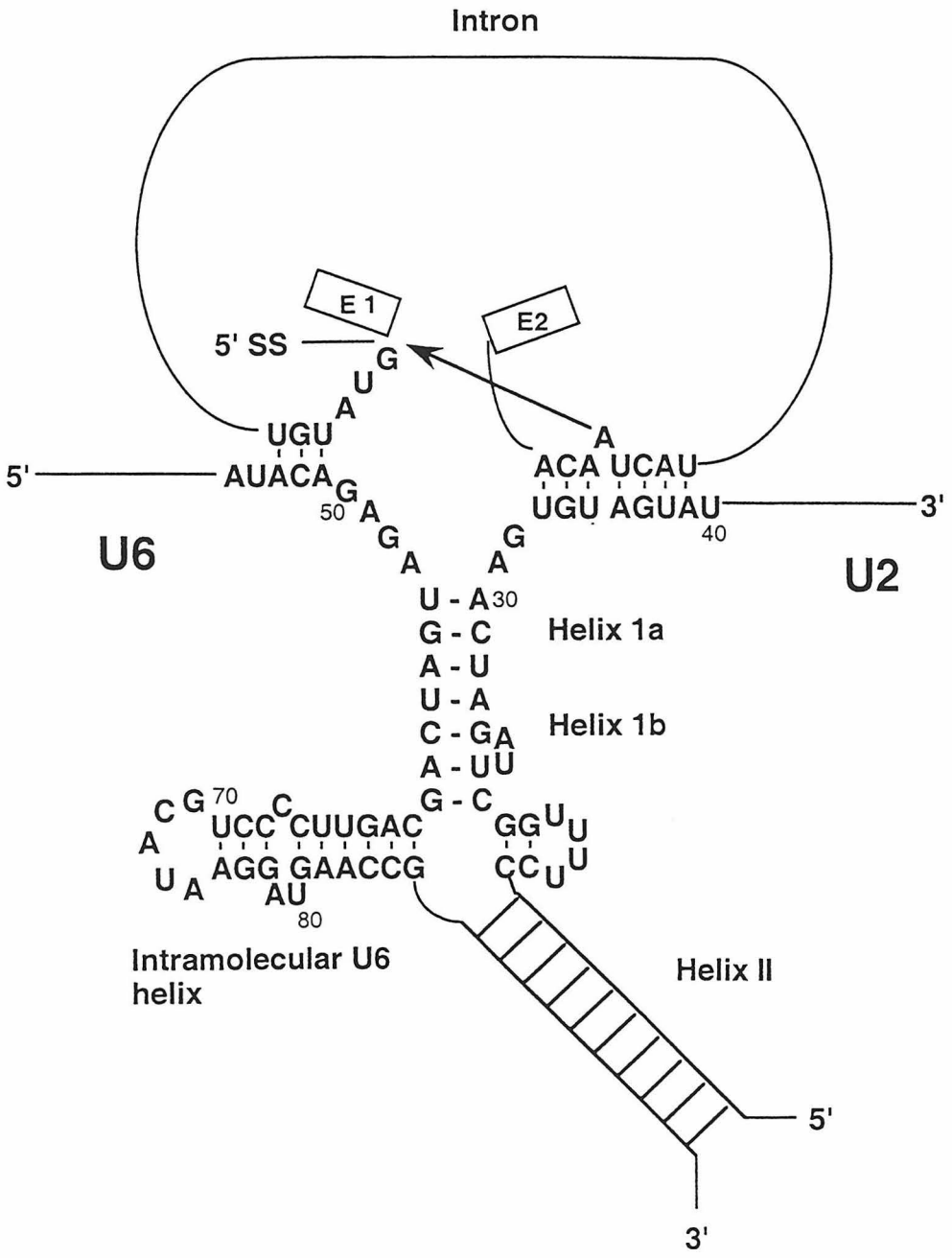
Figure 1A. Spliceosome assembly pathway based on results from yeast and mammalian systems. The abbreviations CC, B, A2-1, A1, A2-2 and A2-3 represent ribonucleoprotein complexes that have been characterized biochemically or genetically. The names of A type complexes and precursor RNA processing (Prp) protein factors are as defined in yeast. In each distinct complex, many protein factors bind to the complex, but have been omitted for simplicity or lack of data; only the Prp proteins which are required for the progression from one complex to another are shown where applicable. Transitions requiring ATP are indicated. The five small nuclear ribonucleoproteins (snRNPs) -- U1, U2, U4, U5 and U6 are enclosed and highlighted in circles. Base-pairing or other types of contacts suggested by genetic and biochemical experiments between the snRNAs and the splicing substrate or other snRNAs are indicated by a series of short bars.


Spliceosome assembly
U1
exon 1
exon 2
Commitment Complex
(CC)
Prp21
Prp5 Prp11
Prp9
U2
ATP
U1
U2
exon 1
exon 2
Pre-Spliceosome
(B)
U2
exon 1 exon 2
U5
U6
ATP
Prp22
U5
U6
U4
ATP
Prp4
Prp24
U6
U4
U1
U2
exon 1
exon 2
A2-1
U5
U6
U4
Second step
U6
U2
exon 1
exon 2
A2-3
U5
U4
ATP
U6
U2
exon 1
exon 2
A2-3
U5
U6
U2
exon 1
exon 2
A1
U5
ATP
A2-2
Prp16
Prp17
Prp18
Slu7
U6
U2
exon 1
exon 2
U5
ATP
Prp2
First step

essential. In the next step, the U2 snRNP joins this complex to form a stable complex called the "pre-spliceosome". In this step a sequence in U2 near its 5' end base-pairs with the conserved intron branch point sequence near the 3' splice site. U4 and U6 are found in a single snRNP containing a U4/U6 RNA duplex involving 21 base pairs in two helices (Guthrie and Patterson, 1988). The U4/U6 snRNP is joined with the U5 snRNP in a triple snRNP which enters the spliceosome in the next step of assembly. In this step a number of important new snRNA/intron and snRNA/snRNA interactions occur. (i) The 5' end of the highly conserved ACAGAGA sequence of U6 base pairs with the 5' splice site, replacing U1/intron interactions and destabilizing interactions between U1 and the spliceosome (Lesser and Guthrie, 1993; Kandels-Lewis and Seraphin, 1993). (ii) Prior to the first step of splicing, three nucleotide bases in a conserved loop of U5 snRNA interact with the last three bases of exon 1 (Newman and Norman, 1992; Sontheimer and Steitz, 1993; Cortes et al.; 1993). The U6 and U5 interactions with the 5' splice site are both crucial for defining the splice site. (iii) A sequence near the 3' end of U6 base pairs with the 5' end of U2 (Datta and Weiner, 1991; Wu and Manley, 1991) to form a helix important for splicing in mammalian cells, but perhaps not in yeast (Fabrizio et al., 1989; Madhani et al., 1990).

In the next step of assembly a crucial and quite remarkable event occurs. U4 either leaves the spliceosome or its interaction is greatly destabilized as both U4/U6 RNA helices dissociate. The presence of U4 is not necessary for either of the catalytic reactions (Yean and Lin, 1991). Dissociation of the U4/U6 helices allows U6 to form a U2/U6 helix and an intramolecular stem-loop near the 3' end of U6 (Madhani and Guthrie, 1992; Fortner et al., 1994; Brow & Vidaver, 1995). These new interactions serve to hold the pre-mRNA so that the 5' splice site is in close proximity with the branch point sequence (Fig. 1B). Presumably at this point the catalytic center of the spliceosome for the first step is nearly complete. However, one further ATP-dependent step, mediated by yeast protein Prp2, is required to activate the first step of splicing. In this step a marked conformational change occurs in the spliceosome, but this has not yet been defined (Kim and Lin, 1993).

Figure 1B. RNA interactions in the fully assembled spliceosome for the first chemical step (adapted from Yan and Ares, 1996). The intermolecular pairing between U6 and U2 and with the pre-mRNA is thought to position the 2'-OH of the branchpoint A (shadowed) close to the 5' splice site. Dashes indicate base-paired regions demonstrated by compensatory mutations. The 5' and 3' exons are indicated as E1 and E2, respectively. 5' SS refers to 5' splice site.


Intron
E 1
E2
5' SS
G
U
A
UGU
A
5'
AUACA
ACA UCAU
UGU AGUAU
3'
50
40
U6
U2
G
A
A
U - A30
G - C
Helix 1a
A - U
U - A
C - GA
Helix 1b
A - UU
G - C
C G 70
C
UCC CUUGAC
GG U U
A
U A
AGG GAACCG
CC U U
AU
80
Intramolecular U6
helix
Helix II
5'
3'

Activation of the spliceosome for the second step of splicing requires a second ATP-dependent step and presumably another conformational change. At least four yeast proteins are required for this step (reviewed by Umen and Guthrie, 1995). Before the second step of splicing takes place, the conserved loop of U5 interacts with the first two bases of exon 2 (Newman and Norman, 1992). The U5 interactions serve to define the 3' splice site and are essential for the second step. Thus it appears that the active site shifts between the two steps, an assumption supported by the fact that the stereochemical requirements (Sp phosphorothioate diastereomer) for the phosphodiester transesterification reaction are identical for the first and second step despite the inversion of configuration after the first step (Moore and Sharp, 1993).

The spliceosome assembly pathway conceptualized in Figure 1A is based on consistent results from the yeast and mammalian systems. Although it is a good model, it is almost certainly oversimplified. The particular order of base-pairing events concomitant with and following the addition of the triple snRNP needs to be further defined. Moreover the roles of proteins in the assembly steps are poorly defined and need further study. To enable these studies, we are undertaking a comprehensive program to more precisely define the RNA-RNA and RNA-protein interactions that take place in spliceosome assembly and to determine at which stages of assembly these interactions take place. The experimental methods for this program are as follows: 1) we have developed an in vitro reconstitution method in which U6 or U2 snRNA in a yeast extract are inactivated via oligonucleotide directed RNase H digestion (Fabrizio et al., 1989; McPheeters et al., 1989). Splicing activity is efficiently restored in the extract by the addition of exogenous synthetic snRNA. 2) In this paper we show that it is possible to incorporate a photo-active crosslinking reagent, 4-thiouridine (4-thioU) at specific positions in U6 RNA. The modified RNA is fully active in reconstitution and in some cases highly efficient crosslinks were observed to form between U6 and the pre-mRNA under conditions of spliceosome formation. 3) To map the stage that these crosslinks occur we plan to employ extracts from temperature

sensitive yeast mutants defective in pre-mRNA splicing. We have shown that these extracts can be inactivated by heat treatment and that activity can be restored by addition of the purified missing protein. The inactivated mutant extracts are blocked in spliceosome assembly at specific stages, so that determining if a particular crosslink can form in a mutant extract gives one piece of information on the stage of that interaction in spliceosome assembly. By carrying out the crosslinking experiments in a number of extracts the stage at which the interaction first occurs, and when it may disappear, can be determined. Once the stage a particular crosslink occurs is determined, that crosslink represents a valuable marker for that stage. An increasing number of yeast mutants, defective in pre-mRNA splicing, are being isolated and should provide an ever more detailed map of the assembly process.

In this paper we demonstrate that every stage of this program can be done. 4-thioU was incorporated at each of four distinct positions in U6 RNA in a region known to form crucial interactions with other spliceosomal components. These RNAs were active in the reconstitution of splicing activity in U6-deficient extracts. Two of the RNAs, 4-thioU46 and 4-thioU51 formed UV-dependent, efficient crosslinks with nucleotides near the 5' splice sites. Crosslinks were also formed in prp2-1 heat inactivated extracts. These extracts form spliceosomes but are blocked just prior to the first chemical step. In these extracts, the 4-thioU U6 RNA formed crosslinks with the pre-mRNA. Addition of purified Prp2 protein to the extract allowed crosslinks to form with the lariat-exon 2 intermediate. This information establishes that the interaction of U6 RNA with nucleotides near the 5' splice site occurs at a stage prior to the first step of splicing and that the interaction continues to be present after the first step. More mutant extracts will have to be tested to determine when the interaction first takes place. Significantly, we have demonstrated for the first time that the crosslink between U6 and the pre-mRNA at the 5' splice site is functional since this crosslink formed in prp2-1 extracts can subsequently be chased into a U6/lariat intermediate product upon complementation with Prp2 protein.

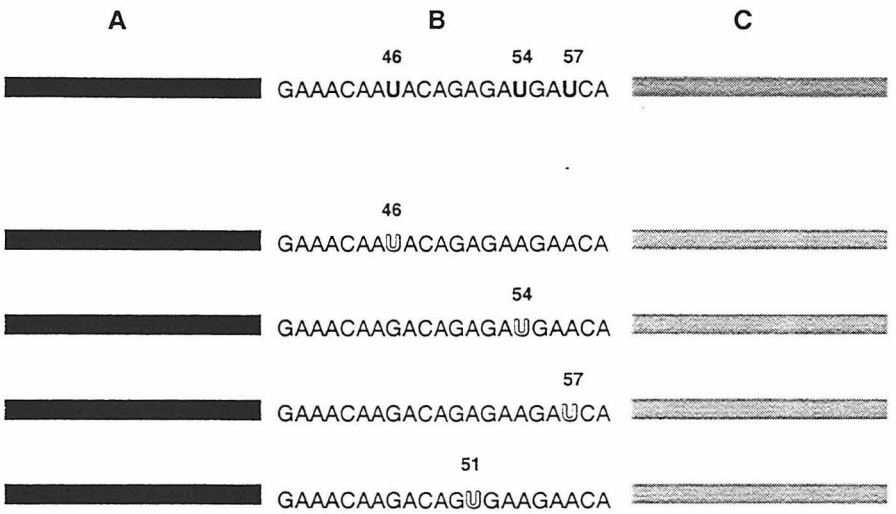
Results

Synthesis of U6 RNA containing site specific 4-thiouridine modifications

Our strategy for this initial step in the investigation of RNA-RNA crosslinks during spliceosome assembly was first to investigate U6 RNA interactions. U6 is the smallest and most conserved of the spliceosomal RNAs (Guthrie and Patterson, 1988) and the reconstitution assay for U6 is efficient. In a comprehensive survey of the effects of U6 mutants, both in vitro (Fabrizio and Abelson, 1990; Fabrizio and Abelson, 1992) and in vivo (Madhani et al., 1990), most of the nucleotides crucial for U6 function were found in a central region of U6 between positions 46 and 61. We therefore sought to develop a system which would allow us to insert specific crosslinking agents in that region. In the region between nucleotides 39 and 59 there are only three U residues at positions 46, 54 and 57. Changes at 46 and 57 produced little or no effect in the reconstitution assay and changes in 54 had not been tested. Our strategy was therefore to synthesize U6 RNA in three pieces A, B and C (Figure 2) and to ligate them together using the template directed ligation technique of Moore and Sharp (1992). Nucleotide substitutions in the central fragment at positions 46, 54 and 57 allowed synthesis of fragment B variants containing a single uridine. Transcription of this fragment by T7 RNA polymerase, with the substitution of 4-thioUTP for UTP, allowed the incorporation of a single 4-thioU into this fragment. So, for example, to investigate the effects of 4-thioU at position 46, U54 and U57 were changed to A. To incorporate 4-thioU specifically at other positions, for example at position 51, a quadruple mutant was necessary; U46 was substituted with a G, U54 and U57 were changed to A and A51 was changed to U.

Fragment B was internally ^{32}P labeled by transcribing with [α–^{32}P] CTP; A and C were not radioactively labeled. The 5'-end of the transcripts B and C were

Figure 2. Site-specific incorporation of 4-thiouridine (4-thioU) into U6 snRNA. The wild-type sequence in the central fragment B of U6 has only 3 uridines (indicated in bold). We incorporated a single 4-thioU (shadowed) into B by transcribing from a synthetic DNA template mutated to encode a single U with 4-thioUTP substituting for UTP. A DNA template assisted 3-way ligation of B with A and C using T4 DNA ligase (Moore and Sharp, 1992) yielded the full length U6.


A
B
C
46
54
57
GAAACAAUACAGAGAUGAUCA
46
GAAACAAUACAGAGAAGAACA
54
GAAACAAGACAGAGAUGAACA
57
GAAACAAGACAGAGAAGAUCA
51
GAAACAAGACAGUGAAGAACA

monophosphorylated by addition of a four-fold excess of guanosine monophosphate to the transcription reaction. This provides the required phosphate group donor at the ligation junction. However, initiation by GMP is not quantitative and accounts for losses in the final ligation yield. Moreover, the heterogeneity at the 3'-end (the N+1 adduct to the transcript) reduces the overall ligation efficiency. The N+1 band does not ligate because it does not base-pair perfectly to the template at the ligation junction. Despite these problems, the three fragments can be ligated with an efficiency of ~30%. From a typical ligation reaction 10 pmoles of ligated U6, a sufficient quantity to carry out 100 reconstitution experiments could be obtained.

U6 snRNAs containing 4-thiouridine can functionally reconstitute splicing activity

In order to inactivate the endogenous U6 snRNA, yeast splicing extracts were incubated for 30 min at 37°C with the oligonucleotide d1 which is complementary to nucleotides 28 to 54 in U6 (Fabrizio et al., 1989). The endogenous RNaseH activity in the extract degrades U6 RNA completely under these conditions and the splicing activity of the extract is lost (Fig. 3 and 4, lane 2). During the 30 min incubation, the d1 oligonucleotide itself is degraded by the DNases in the extract, so that the complementing U6 RNA added later is not cleaved and splicing activity is restored (Fig. 3 and 4, lane 3). Usually, a large batch of extract was treated with d1, and aliquots were taken to test U6 reconstitution and crosslinking reactions under various conditions. We have found that to obtain reproducible reconstitution results it is necessary to carefully determine the optimal concentration of oligonucleotide d1 to add to each extract. Apparently there is a precarious balance between the amount of oligonucleotide necessary to achieve complete cleavage of endogenous U6 RNA and the ability of DNases in the extract to degrade d1, allowing the survival of the complementing synthetic RNA.

Figure 3. Reconstitution of splicing activity and crosslinking of ^{32}P-labeled actin pre-mRNA with U6 containing 4-thioU at position 51. The positions of pre-mRNA substrate, lariat-exon 2 (IVS-E2) intermediate, and spliced product (mRNA) are indicated. (Lariat intron product cannot be seen because it co-migrated with pre-mRNA on this gel.) with Yeast splicing extract was first treated with an oligonucleotide d1 complementary to a U6 sequence to cleave the endogenous U6 RNA by RNAse H activity of the extract (lane 2). Both steps of splicing were restored by the addition of synthetic U6 RNA transcript without any modification (lane 3) while only the first step is restored by the ligated U6 containing 4-thioU at position 51 (lane 4). Splicing activity is not restored in the absence of ATP (lane 5). Irradiation with UV (325nm) (lane 6, 7, and 8) reproducibly results in three 4-thioU (cf. lane 6 and 8) and ATP-dependent (cf. lane 7 and 8) crosslinks: crosslink 1, 2 and 3 (lane 8). All reactions were de-proteinized and fractionated on an 8 % denaturing polyacrylamide gel.

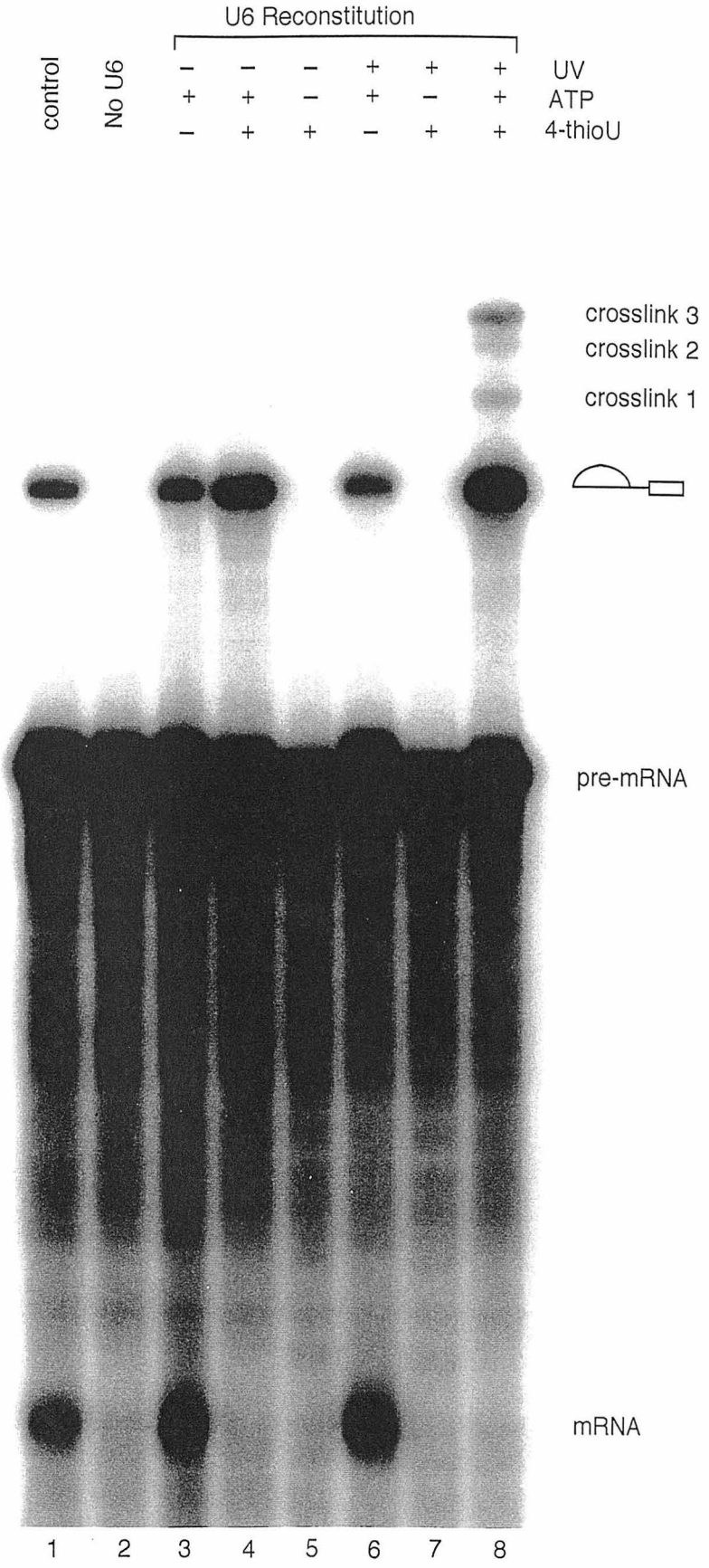
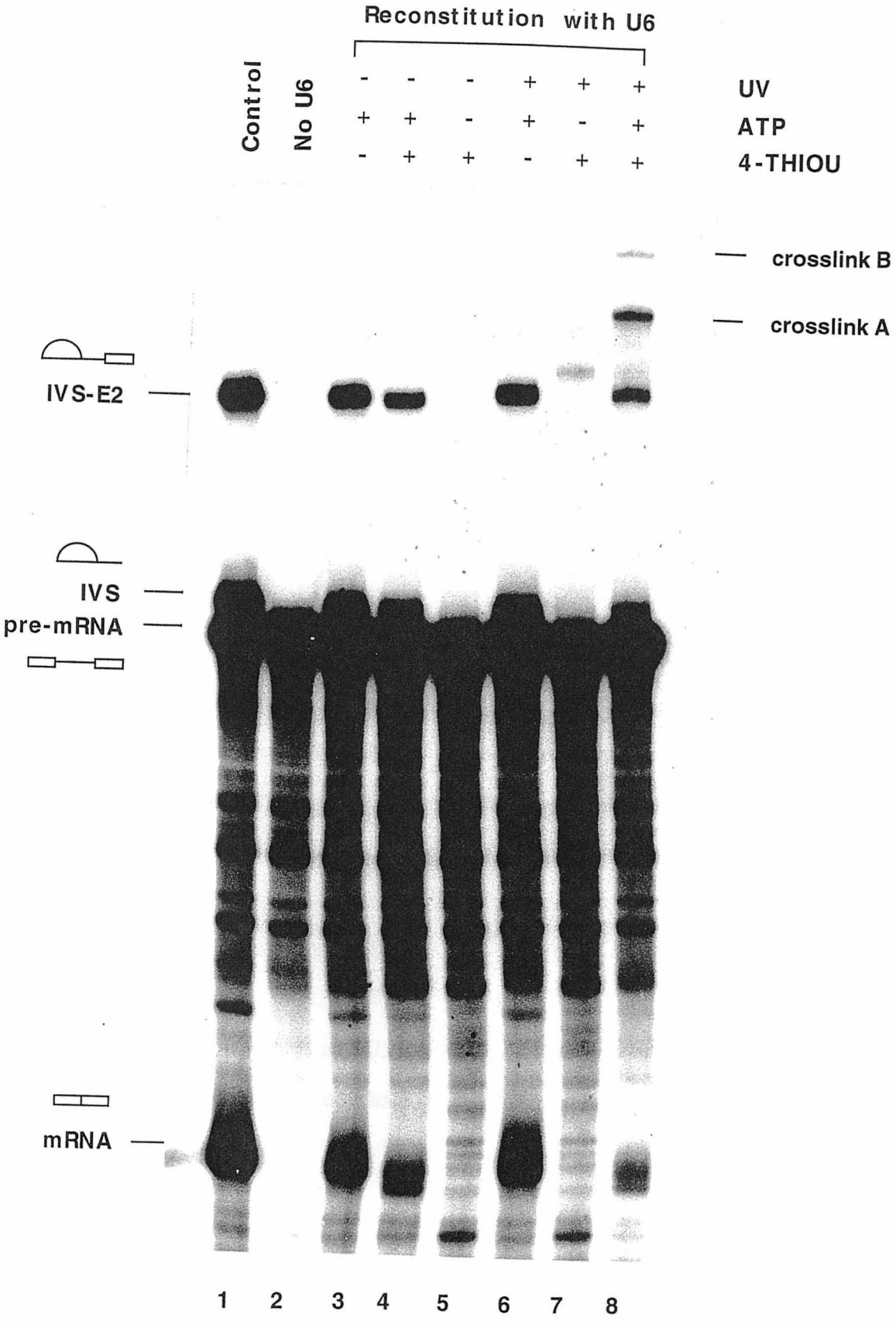

U6 Reconstitution
control
No U6
UV
ATP
4-thioU
crosslink 3
crosslink 2
crosslink 1
pre-mRNA
mRNA
1
2
3
4
5
6
7
8

Figure 4. Reconstitution of splicing activity and crosslinking of ^{32}-P labeled actin pre-mRNA with U6 containing 4-thioU at position 46. The experimental set-up and order of lanes is as in Fig. 3. However, splicing activities of both the first and the second steps were reconstituted by the addition of ligated U6 with 4-thioU at position 46 (lane 4). Irradiation with UV (325nm) (lanes 6, 7 and 8) reproducibly results in two 4-thioU and ATP-dependent crosslinks: crosslink A (high yield -- 20% of spliced substrate) and a crosslink B (minor band).


Reconstitution with U6
Control
No U6
- - - + + +
+ + - + - +
- + + - + +
UV
ATP
4-THIOU
crosslink B
crosslink A
IVS-E2
IVS
pre-mRNA
mRNA
1 2 3 4 5 6 7 8

Lane 1 of Figures 3 and 4 shows the control splicing reaction of an actin pre-mRNA substrate in yeast splicing extract. On a 5% gel, the lariat intron product (without the 3' exon) can sometimes co-migrate with the pre-mRNA (cf. lane 1 of Fig. 3 and Fig. 4). Lanes 2, 3, and 4 in Figures 3 and 4 show the complete loss of splicing activity due to incubation with oligonucleotide d1, and the reconstitution of splicing activity in these depleted extracts by a U6 wild-type transcript (lane 3) and 4-thioU modified U6 RNA (lane 4) containing a quadruple (Fig. 3) or double mutation (Fig. 4). In all the splicing and crosslinking experiments, the pre-mRNA was internally labeled with high specific activity [α -^{32}P]-UTP. The ligated U6 was also internally labeled (as discussed above) but to a much lower specific activity than the pre-mRNA. Due to the high specific activity internal labeling of the 195 uridines in the actin pre-mRNA sequence, many non-specific degradation fragments can be observed.

All synthetic U6 RNAs containing the 4-thioU modification and the double mutations were capable of functionally reconstituting splicing activity in extracts depleted of endogenous U6 RNA with an efficiency similar to that of synthetic wild-type U6 RNA (cf. lanes 2 and 4 in Fig. 3 and Fig. 4 for reconstitution by synthetic U6 modified with 4-thioU at positions 51 and 46; data not shown for U6 containing 4-thioUs at positions 54 and 57). The following three double mutations in U6 RNA did not interfere with the reconstitution of splicing activity : 1) U54 to A54 and U57 to A57, 2) U46 to G46 and U57 to A57, and 3) U46 to G46 and U54 to A 54. The quadruple mutant containing 4-thioU at position 51, U46 to G, U54 to A and U57 to A faithfully reproduced the in vitro phenotype of second step block and lariat-exon2 accumulation characteristic of all single base changes at the invariant A51 adenosine (Fig. 3, lane 4) (Fabrizio and Abelson, 1990).

U6 snRNA containing specific 4-thioU residues can be crosslinked to pre-mRNA and to the lariat-exon2 intermediate

4-Thiouridine is an analog of uridine with a sulfur atom replacing the keto oxygen at the 4 position in the pyrimidine ring. Upon photoexcitation to its triplet state it can form covalent bonds with either bases of a nucleic acid or the amino acids of a protein in close proximity. The photochemical mechanism is thought to be vis a vis (2 + 2) cycloadditions with double bonds or coupling reactions with hydrogen donors via free radical intermediates (Favre et al., 1991). 4-thioU forms most efficient RNA-RNA crosslinks when it is in a stacked conformation rather than base-paired (Favre, 1990). Its strongest advantage is that it is photoactivated by low energy (>300nm) UV light which does not activate unmodified proteins or other nucleic acids in the spliceosome, unlike the high energy 254 nm UV light used in conventional UV crosslinking. This means that the spliceosome is more likely to be undamaged when irradiated with 300 nm UV light since photodamage is minimized and also, nonspecific crosslinks are minimized. For an excellent and thorough methodology paper on site-specific RNA crosslinking with 4-thiouridine refer to Sontheimer, 1994.

Irradiation of extracts reconstituted by 4-thioU51 U6 RNA with a 10mW UV LASER (Liconix) at 325nm for 5 min produced three slowly migrating species (denoted as crosslinks 1, 2, and 3, Figure 3). The appearance of these species is dependent on UV, on the addition of pre-mRNA and ATP and are not seen when the complementing RNA does not contain 4-thioU (cf. lanes 3-7 with 8, Figure 3). The dependence of the crosslinks on ATP and pre-mRNA suggest that they are relevant to functional spliceosome formation. In the absence of ATP, spliceosome assembly does not occur. The 4-thioU dependence indicates that the site of the crosslink on U6 is the 4-thioU modified site (in this case the 4-thioU51) because the 4-thioU modification is not expected to photoactivate other sites (Wyatt et al., 1992; Sontheimer, 1994; Newman et al., 1995). The cumulative total of the

crosslink yields are about 30% of the spliced lariat intermediate as measured by densitometry. The high yields obtained for the crosslinks with 4-thioU51 in U6 can be attributed to an accumulation of spliceosome complexes due to the second step block phenotype caused by the mutation at A51. Of the three crosslinked species, crosslink 3 is three times higher in yield than either of the other two crosslinks as quantitated by densitometry on Molecular Dynamics phosphorimager. Crosslink 1 and 2 have approximately comparable yields. Higher yields of the crosslinks were obtained by increasing the concentration of U6 4-thioU51 up to the point of saturation (data not shown). Typically, 50 to 100 fmoles of 4-thioU labeled U6 in a 5 μl reaction gave the yield shown in Figure 3. The crosslinks obtained with 4-thioU51 U6 could be repeated in different batches of wild-type splicing extracts, ligated U6 and pre-mRNA splicing substrate (data not shown). Higher crosslink yields were obtained in extracts that gave more efficient U6 reconstitution (data not shown).

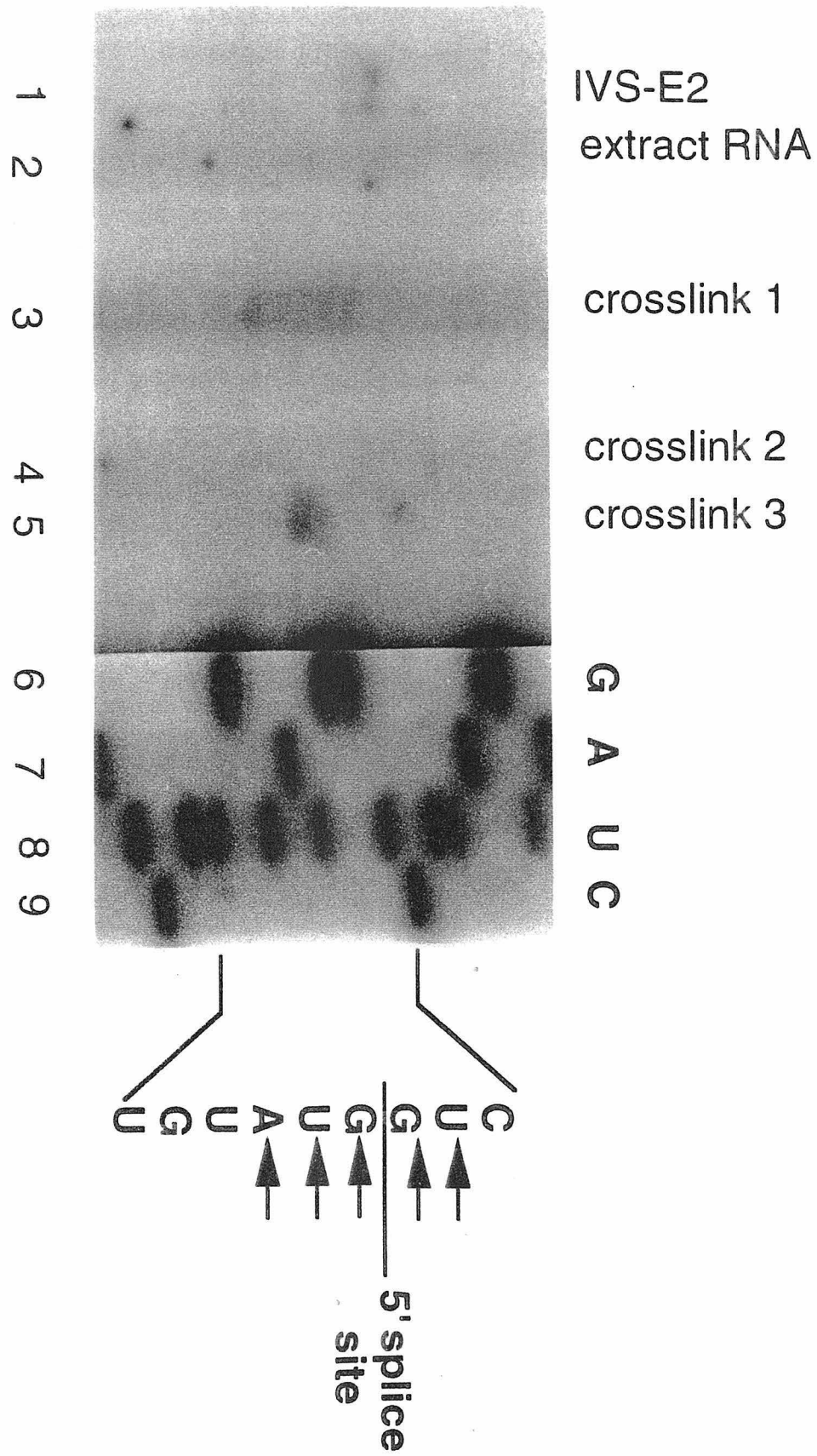
Two crosslinks (A and B) were also observed when extracts complemented by 4-thioU46 U6 RNA were irradiated (lane 8, Fig. 4). Again the appearance of the crosslinks was dependent on the addition of ATP and pre-mRNA, and on a 4-thioU complementing RNA. The mobility of the 4-thioU46 crosslink A was similar to that of the 4-thioU51 crosslink 1 in Figure 3, but in lower yields, analogous bands to 4-thioU51 crosslinks 2 and 3 (i.e. crosslink B) were also observed in this case. Two ATP-independent crosslinks of faster mobility (lane 7, Fig. 4) were obtained in much lower yields and were not reproducible in other experiments. No attempt was made to characterize them. Crosslinks were not observed when U6 RNA containing 4-thioU at positions 54 and 57 were irradiated though both of these RNAs efficiently reconstituted oligonucleotide d1-inactivated extracts (data not shown).

4-thioU at position 51 crosslinks to sequences in the intron near the 5'splice site

Preliminary experiments established that, as expected, the 4-thioU51 crosslink is between U6 RNA and the pre-mRNA or reaction intermediates. RNAseH digestion, directed by oligonucleotides complementary to sequences in the 5' exon, intron and 3' exon, showed that crosslink 1 was susceptible to cleavage by all three probes, while crosslinks 2 and 3 were digested only with probes to the intron and the 3' exon (data not shown). This result suggested that crosslink 1 is to the pre-mRNA, and that crosslinks 2 and 3 are to the lariat intermediate. Primer extensions with a primer complementary to sequences in exon 2 firmly established that crosslink 1 is a crosslink to the pre-mRNA because the primer extended past the branched structure all the way to the site of the crosslink near the 5' splice site, while crosslink 2 and 3 were shown to be to the lariat intermediate because primer extensions were blocked at the lariat branched nucleotide (data not shown).

The site of the crosslink to the pre-mRNA (for crosslink 1, in Fig. 3) or to the lariat intermediates (crosslinks 2 and 3, Fig. 3) were mapped to nucleotide resolution by identifying the site where the crosslink blocks primer extension by reverse transcriptase. The site of the crosslink is 1 nucleotide 5' to the site where primer extension is blocked (Newman et al., 1995). A primer complementary to nucleotides 39-55 in the intron was used to prime DNA synthesis by reverse transcriptase on the gel-purified crosslinked RNA template. 0.1 fmoles of crosslinked RNA were required to identify the crosslink site by this method. This amount could routinely be obtained by isolating the crosslinked species from five 5μl reaction mixtures. The site of primer extension blockage was absolutely assigned by comparing the position of the stops to a dideoxy sequencing ladder (Fig. 5, lanes 6-9) of the actin gene plasmid template using the same primer. For crosslink 1, primer extension was stopped in five positions (Fig. 5, lane 3). Thus, the site of crosslink

Figure 5. Primer extension mapping of sites on pre-mRNA which form a crosslink with 4-thioU51 in yeast U6 RNA. The three crosslinks 1, 2 and 3 (lanes 3-5) from Fig. 3 were purified and used as templates for reverse transcriptase primed by a 5'-end labeled deoxynucleotide Int complementary to sequences in the intron (positions +39 to +55) near the 5' splice site. UV-irradiated lariat intermediate (lane 1) was used as a positive control for the accuracy of the reverse transcription reaction and bands of gel-fractionated extract RNA co-migrating with the crosslinks were used to rule out the possibility that the extension stops are due to fortuitously co-migrating RNAs in the extract. Lanes 6-7: Dideoxy sequencing lanes to identify the sites of primer extension blockage. Products of all reactions were fractionated on denaturing 8% polyacrylamide gels. Sites of crosslinking (one nt 5' to the positions of primer extension blockage) are indicated by unlabeled dark arrows next to the sequence.


1 2 3 4 5 6 7 8 9
IVS-E2
extract RNA
crosslink 1
crosslink 2
crosslink 3
G A U C
C U G G U U A U G U
5' splice site

was determined to be at positions -2, -1, +1, +2, +3, (UG/GUA) at the 5' splice site (see arrows on the side of the sequencing lanes). Densitometry measurements on a Molecular Dynamics Phosphorimager of this gel showed that the 5 extension stops are equal in strength.

Crosslink 3 (in the lariat exon2 intermediate) was mapped by primer extension to the invariant G_1 at the 5' splice site/intron junction (lane 5, Fig. 5). By comparing the position of these stops to the reverse transcriptase stop due to the branched nucleotide in the lariat intron intermediate (cf lanes 1 and 5, Fig. 5) it can be seen that crosslink 3 is to a position one nucleotide 3' of the branched nucleotide in the lariat intermediate. This crosslink is the reverse of the crosslink found in mammalian extracts between 4-thioU uniquely replacing U in the conserved GU dinucleotide and U6 RNA. In this case the crosslink was between the lariat intermediate and the invariant A45 in the U6 ACAGAGA conserved motif (the yeast equivalent is A51) (Sontheimer and Steitz, 1993). In contrast to the five primer extension stops obtained with crosslink 1, there are only two stops for this crosslink. Sontheimer and Steitz also observed only two stops due to the crosslink between 4-thioU at position 2 in the intron and U6 RNA.

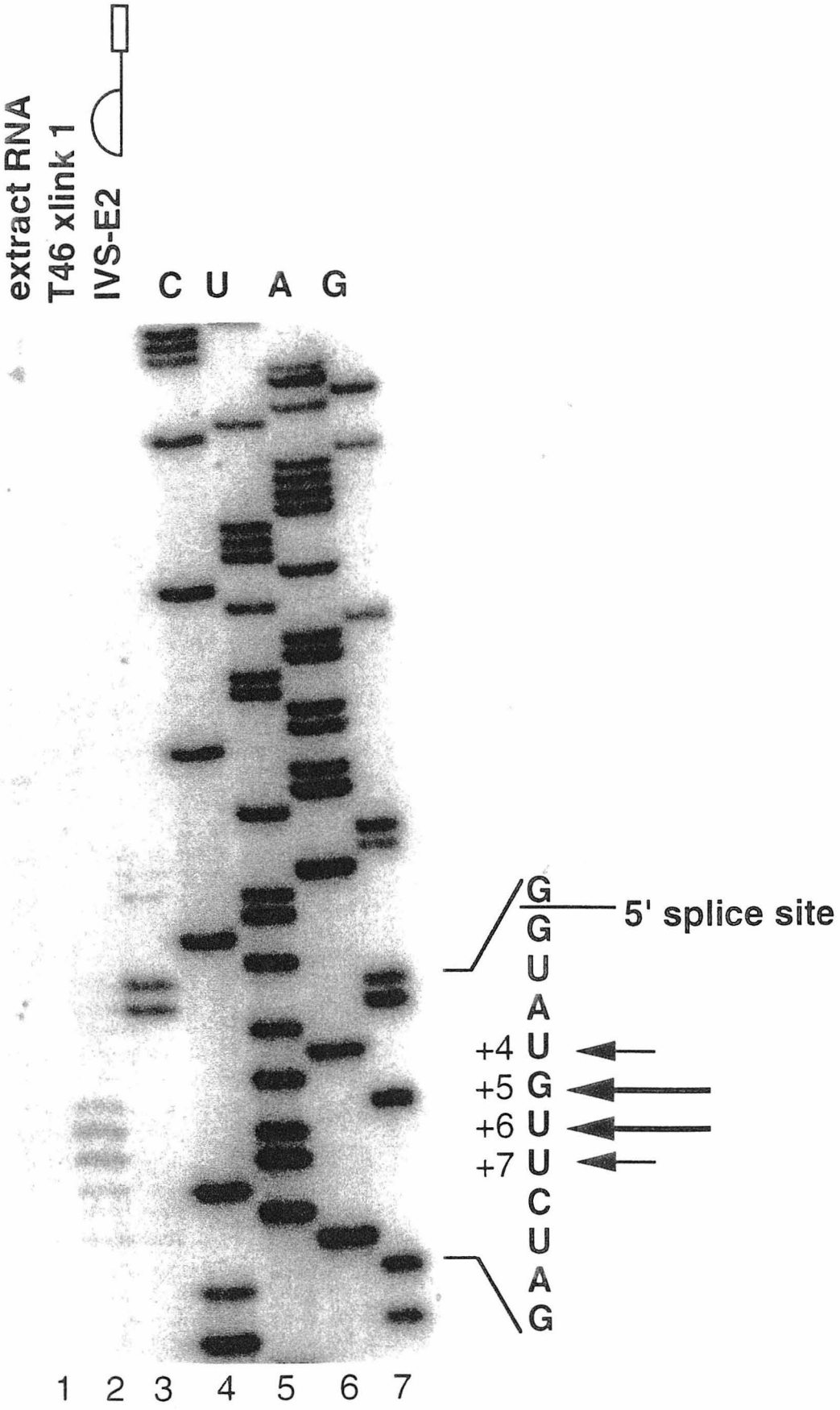
The RNA template from crosslink 2 did not result in any primer extension despite several attempts possibly because the sample yield after elution from gel fragments was too low. Further experiments will be required to determine the site of this crosslink.

For a control of the primer extension reactions, an extension was done on a UV irradiated sample of pre-mRNA (not shown) and on the lariat intermediate (lane 1, Fig. 5) to make sure that the primer extension blocks are not due to a secondary structure or to UV generated nicks in the pre-mRNA or the lariat intermediate. In lane 2, primer extension was performed on UV irradiated extract containing no pre-mRNA, to make sure that the primer extensions attributed to crosslinks were not due to unfortuitous priming of DNA synthesis from some RNA in the extract. No specific bands were seen.

4-ThioU 46 in U6 crosslinks to a site in the intron that is several nucleotides downstream of where the 4-thioU51 U6 -- 5' splice site crosslink occurs

Band A (lane 8, Fig. 4) was determined to be a U6-pre-mRNA crosslink by its sensitivity to RNAseH digestion using probes complementary to seqeuences in the pre-mRNA (data not shown). This was the only product recovered in sufficient quantity for mapping by primer extension. As in the cases of the 4-thioU51 crosslink to the pre-mRNA, reverse transcription with a primer complementary to the 3'exon extended past the branchpoint A, further suggesting that the crosslink is to the pre-mRNA (data not shown). The same primer in the intron, complementary to nucleotides 39-55, was used to prime DNA synthesis by reverse transcriptase. As in the case of the crosslink between pre-mRNA and 4-thioU51 in U6, there were several (in this case 4) primer extension stops (Fig. 6, lane 2). From the primer extension stop sites it was determined that 4-thioU46 in U6 crosslinks to positions +4, +5, +6, and +7, in the intron. Note that in comparison to the positions in the intron where 4-thioU51 forms crosslinks, these crosslinks are as much as five nucleotides downstream (Fig. 10). Control primer extensions were done on pre-mRNA, lariat intermediate and extract treated substrates to make sure that the primer extension stops were specific to the crosslinks. The two major stops shown for the lariat intermediate were reproducible (cf. Fig. 5, lane 1) and suggests that reverse trancriptase may be stuttering at the branch site or adding on an extra nucleotide.

Figure 6. Primer extension mapping of sites on pre-mRNA which form a crosslink with 4-thioU46 in yeast U6 RNA. Crosslink A from Fig. 4 which is denoted as T46xlink1 in this figure (lane 2), along with an equivalent region of a lane containing RNA from extract crosslinked in the absence of substrate (lane 1), and UV-irradiated lariat intermediate (lane 3) were excised from the gel in Fig. 4 and used as RNA templates in a reverse transcription reaction primed by Int (see above). Dideoxynucleotide sequencing ladders of the actin gene plasmid DNA were generated for comparison (lanes 4-7) as before. Sites of crosslinking are indicated by unlabeled dark arrows next to the sequence.

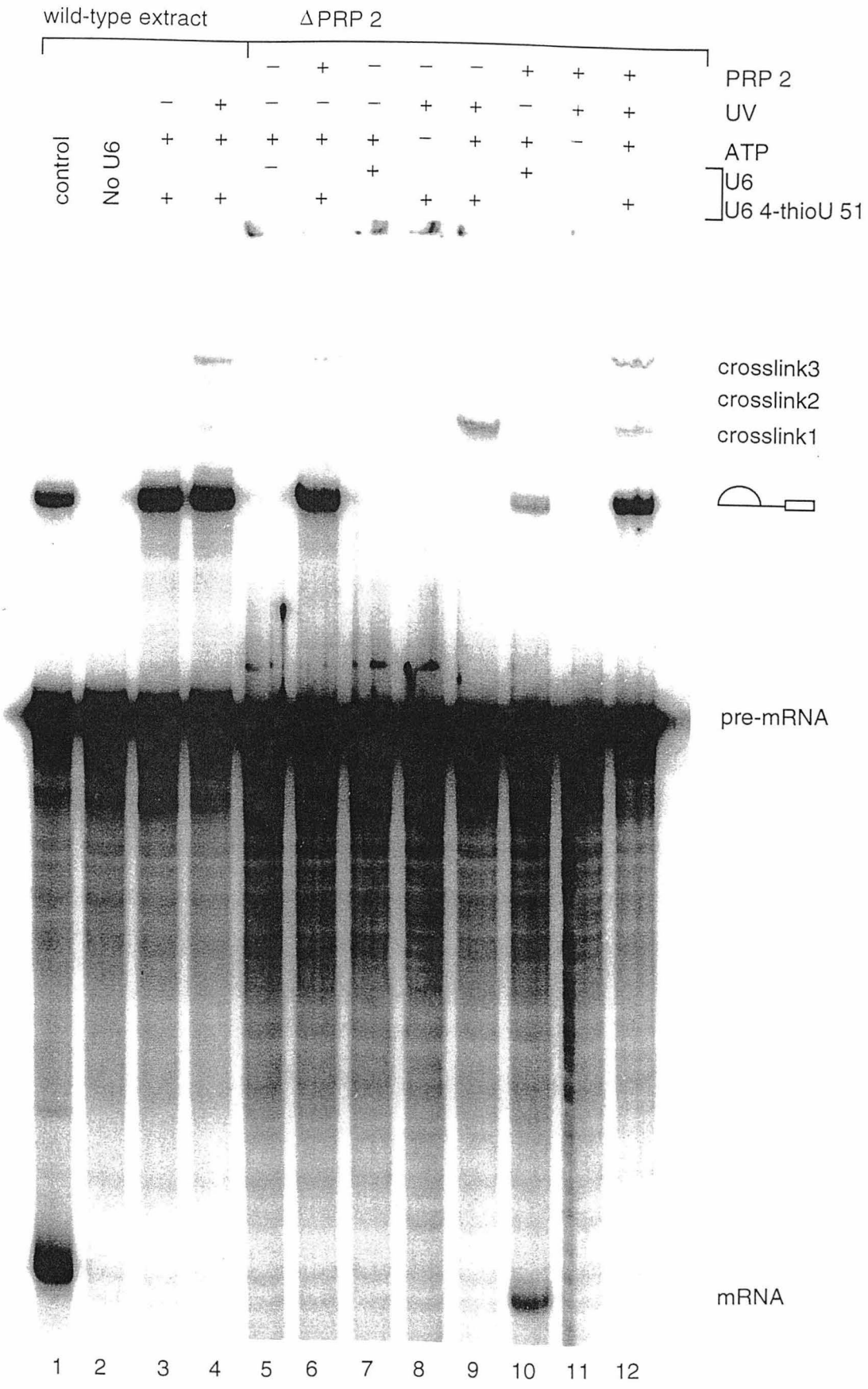

extract RNA
T46 xlink 1
IVS-E2
C U A G
G
G
U
A
+4 U
+5 G
+6 U
+7 U
C
U
A
G
5' splice site
1 2 3 4 5 6 7

The 4-thioU51 and 4-thioU46 crosslinks occur in prp2 blocked spliceosomes

The crosslinks described above demonstrate close interactions between U6 RNA and intron sequences near the 5' splice site. It is crucial to know just when these interactions occur. This information can be obtained by carrying out the crosslinking experiment in extracts that are blocked at different stages of spliceosome assembly. The mutant prp2-1 blocks spliceosome assembly at a stage just prior to the first catalytic step. The spliceosome appears to be fully assembled at this point, and for the first chemical step to occur there is a requirement for the ATP-dependent action of Prp2, a spliceosomal ATPase belonging to the DEAH family, and at least one other factor (Kim and Lin, 1993). We have employed heat inactivated prp2-1 extracts to determine whether the crosslinks to pre-mRNA observed above occur before this stage of assembly.

Splicing extracts derived from the prp2-1 mutant strain were incubated with the d1 oligonucleotide at 37^{o}C to simultaneously degrade the endogenous U6 snRNA and to heat-inactivate the temperature-sensitive prp2-1 protein (lane 5, Fig. 7). Upon addition of exogenous recombinant functional Prp2 protein and synthetic 4-thioU51 U6 RNA, splicing activity is restored (lane 6, Fig. 7). However, due to the second-step block phenotype of the U51 mutant only the first step occurs; the lariat intermediate accumulates strongly and the mRNA product is not observed. When the prp2-1 inactivated extract, reconstituted with 4-thioU51 U6 RNA, is irradiated with UV in the presence of ATP and pre-mRNA, a single crosslink occurs (lane 9, Fig. 7) which co-migrates with the 4-thioU51-pre-mRNA crosslink seen in wild type extracts (cf. crosslink 1 in lane 4 of Fig. 7). The two crosslinks to the lariat intermediate (crosslinks 2 and 3 in lane 4, Fig. 7) observed in wild-type extracts are not observed when the crosslinking experiment is done in a prp2-1 heat inactivated extract (lane 9, Fig. 7). However, these two crosslinks appear when the prp2-1 heat-inactivated extracts are complemented with Prp 2 protein (lane 12, Fig. 6). These

Figure 7. Reconstitution of splicing activity and crosslinking of ^{32}P-labeled actin pre-mRNA with U6 containing 4-thioU at position 51 in prp2 blocked spliceosomes. Lanes 1-4: control crosslinks in the wild-type extract for comparison. Reconstitution and crosslinking in wild-type extract results in three 4-thioU and ATP-dependent crosslinks as before in Fig. 3, lane 8. When extract prepared from a temperature sensitive prp2 mutant yeast strain is incubated with d1 at 37^{o}C for 30 min both endogenous U6 and Prp 2 protein are inactivated (lane 5). Addition of 4-thioU51 U6 and functional Prp 2 protein functionally reconstitutes the first step of splicing (lane 6). The characteristic phenotype of lariat intermediate accumulation is observed. Lanes 7, 8, 9: crosslinking in prp2-1 blocked spliceosomes. The crosslink to the pre-mRNA (crosslink 1) appears in prp2-1 blocked spliceosomes in an ATP and 4-thioU-dependent manner (lane 9). Lanes 10-12: crosslinking in prp2-1 inactivated spliceosomes complemented with a functional Prp2 protein. Addition of Prp 2 protein in the presence of ATP and 4-thioU in U6 results in a reappearance of the additional two crosslinks to the lariat exon-2 intermediate (lane 12).


wild-type extract
ΔPRP 2
PRP 2
UV
ATP
U6
U6 4-thioU 51
control
No U6
crosslink3
crosslink2
crosslink1
pre-mRNA
mRNA
1 2 3 4 5 6 7 8 9 10 11 12

results demonstrate that the crosslink between 4-thioU51 and the 5' splice site in pre-mRNA occurs prior to the action of the Prp 2 gene product.

Similar results were obtained when the 4-thioU46 U6 RNA was used to complement prp2-1 heat inactivated extracts. Crosslink A occurred in prp2-1 inactivated extracts (lane 9, Fig. 8). The low yield crosslink B (to the lariat intermediate) did not form unless Prp2 protein was added to the reaction mixture (lane 12, Fig. 8).

The U6/pre-mRNA crosslinked RNA formed in prp2-1 extracts is a substrate for the first step of splicing-a chase experiment

A key question in any crosslinking experiment is whether one has identified a biochemically relevant or a non-functional interaction. Does the interaction lie on the reaction pathway or is it instead, perhaps, on a discard shunt pathway? Biochemical relevance can be demonstrated if the crosslinked product can be shown, in a chase experiment, to participate in a subsequent step of the reaction. The block to splicing before the first step in prp2-1 heat inactivated extracts allowed us to perform such an experiment. As in the previous experiment heat inactivated prp2-1 extracts were treated with oligonucleotide d1 to inactivate endogenous U6 RNA and were reconstituted by addition of either 4-thioU51 or 4-thioU46 U6 RNA. After incubation for 10 min. the reaction mixture was irradiated on ice for fifteen minutes and then Prp2 protein and an excess of cold pre-mRNA was added and the reaction mixture was incubated for a further 15 min at 23^{o}C (the chase). In Figure 9, (lanes 2, 6) it can be seen that splicing was restored by addition Prp2 protein and in lanes 3, 7 that it was blocked by addition of cold pre-mRNA confirming that no new spliceosomes were formed with labeled pre-mRNA during the chase period. In lanes 4 and 8 it can be seen that only the crosslinked product previously identified as the U6-pre-mRNA crosslink forms in the absence of added Prp2 protein. However, when Prp2 protein is added during the chase period a new crosslink product, previously

Figure 8. Reconstitution of splicing activity and crosslinking of ^{32}P-labeled actin pre-mRNA with U6 containing 4-thioU at position 46 in prp2 blocked spliceosomes. Lanes 1-4: Reconstitution and crosslinking in wild-type extract results in two 4-thioU and ATP-dependent crosslinks as in Fig. 4, lane 8. The double knockout (Prp 2 and U6) results in loss of splicing activity as before (lane 5). Addition of a ligated U6 and functional Prp2 protein functionally reconstitutes both steps of splicing. (lane 6). Lanes 7, 8, 9: crosslinking in prp2-1 blocked spliceosomes. The crosslink to the pre-mRNA (crosslink A) appears in prp2-1 blocked spliceosomes only in the presence of ATP and 4-thioU (lane 9). Lanes 10-12: crosslinking in prp2-1 inactivated spliceosomes complemented with a functional Prp2 protein. Addition of Prp2 protein in the presence of ATP and U6 containing 4-thioU results in a reappearance of the crosslink to the lariat intron-exon2 intermediate (lane 12).

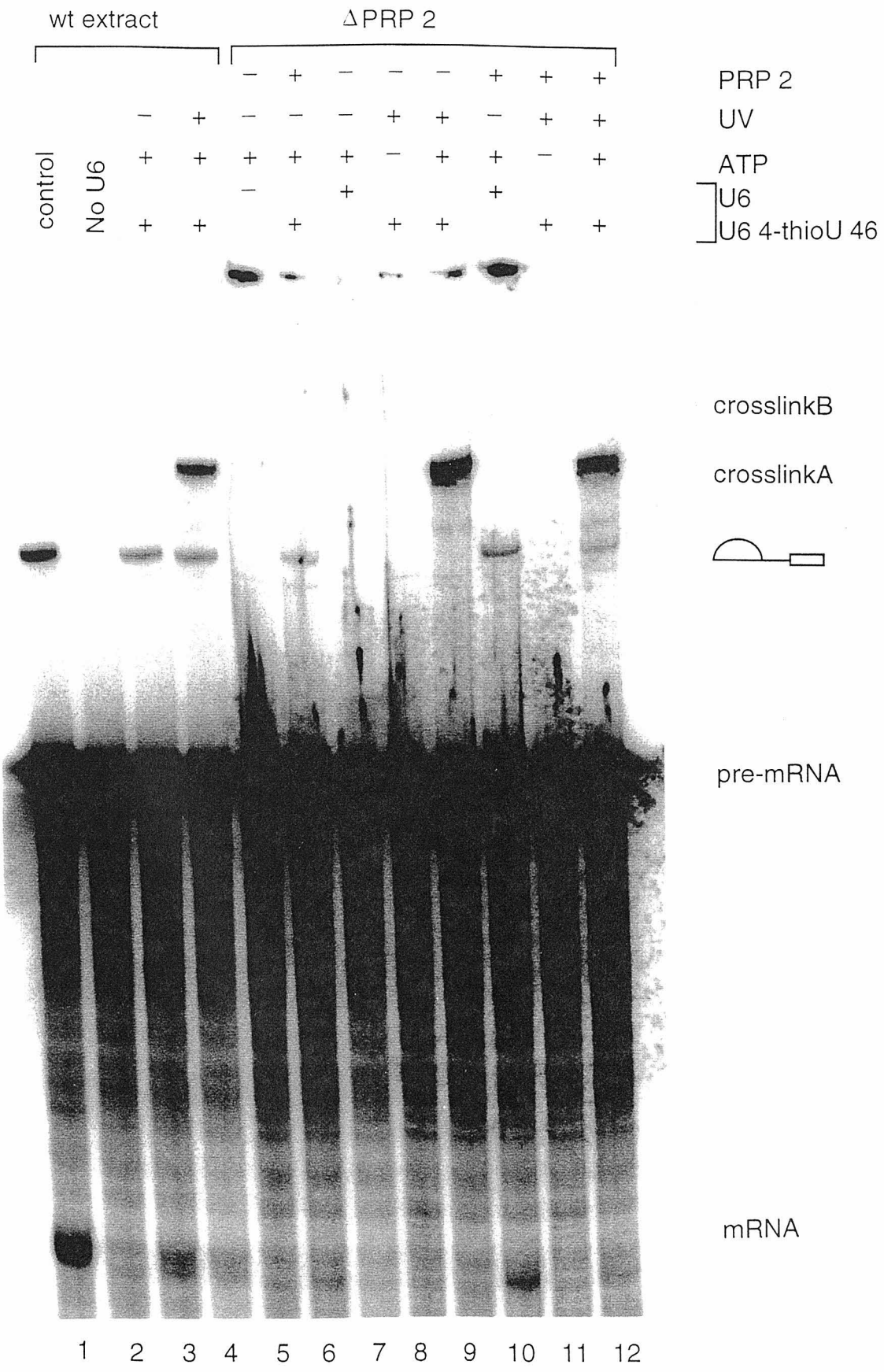
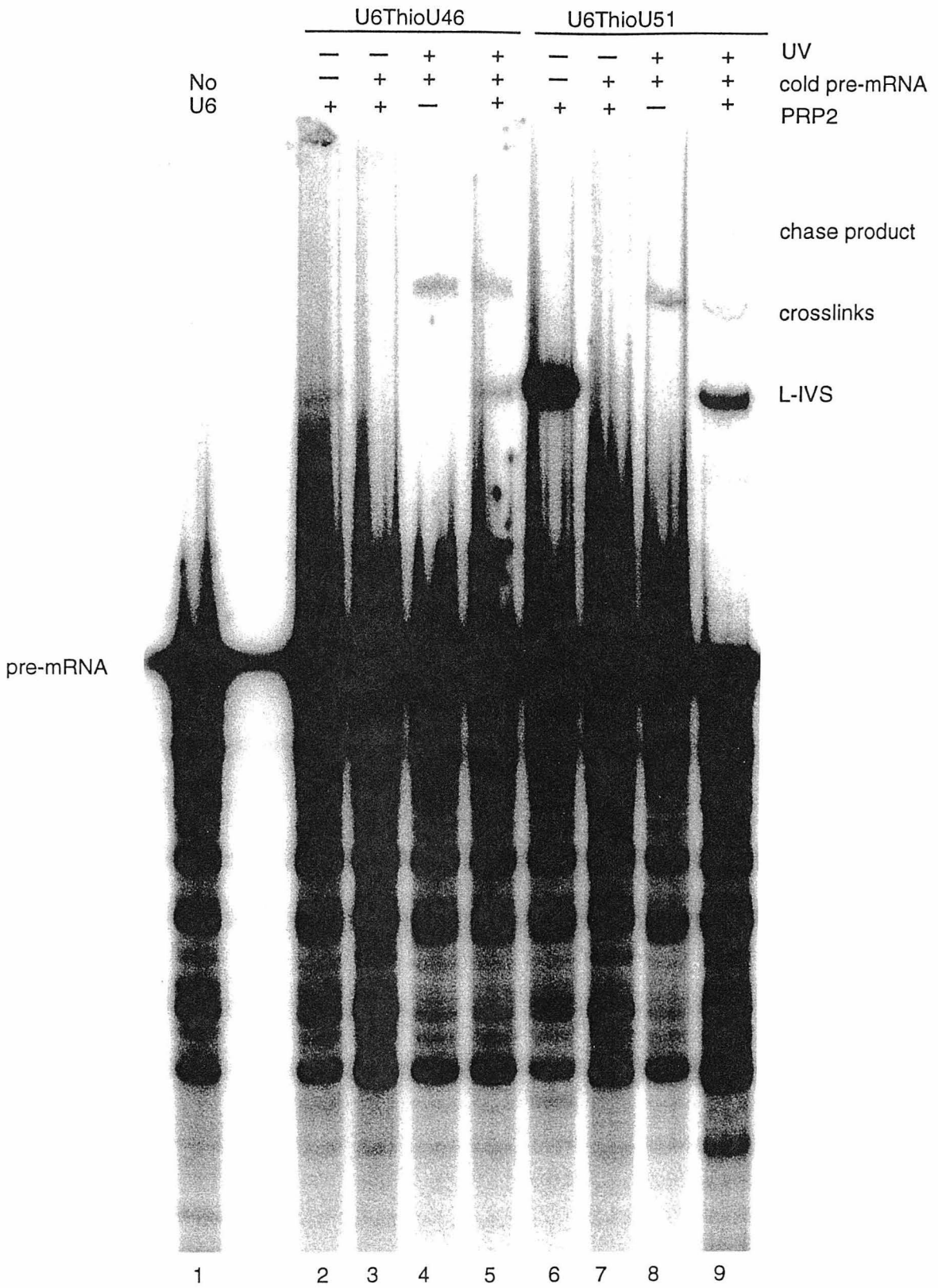

wt extract
ΔPRP 2
PRP 2
UV
ATP
U6
U6 4-thioU 46
control
No U6
crosslinkB
crosslinkA
pre-mRNA
mRNA
1 2 3 4 5 6 7 8 9 10 11 12

Figure 9. Prp2 protein chase of U6/pre-mRNA crosslinks formed in prp2-1 heat inactivated extracts. As in the experiments shown in Fig. 7 and 8, prp2-1 extracts were heat inactivated and depleted of endogenous U6 by incubation with oligonucleotide d1. This treatment abolishes splicing activity (lane1). Upon complementation with Prp2 protein and 4-thioU46 U6, both steps of splicing are restored (lane 2). After incubation for 10 min, the reconstituted extract was UV-irradiated on ice for 15 min. Excess cold pre-mRNA was added (to inhibit assembly of new spliceosomes) and the reaction mixture was incubated for additional 15 min at 23^{o}C in the absence of UV light. The U6 4-thioU46/pre-mRNA crosslink accumulates with high yield in prp2-1 heat inactivated extract, but does not subsequently splice if the extract is not complemented with Prp 2 protein (lane 4). The same crosslink goes onto splice, forming a higher molecular weight band when complemented with Prp2p and incubated further (lane 5). Lane 3 is a control lane which shows that splicing does not proceed despite complementation with both functional Prp2 protein and 4-thioU46 U6 when excess cold pre-mRNA is added. Lanes 6-9: similar experiments as in lanes 2-5, but using U6 4-thioU51. Note the characteristic second step block phenotype and strong accumulation of lariat intermediate in lane 6. The 4-thioU51/pre-mRNA crosslink formed in prp2-1 extract is very efficiently chased into splicing intermediate upon complementation with functional Prp2 protein (lane 9). identified as U6-lariat intron/exon2, is seen. This chase product is in much higher yield in the case of 4-thioU51 (lane 9) than in 4-thioU46 (lane 5) where only a trace of chase product can be seen. This result indicates that the interaction between the nucleotide at position 51 and the pre-mRNA is a functional interaction and that the presence of the crosslink does not deter those complexes from proceeding through the first step of the reaction.


U6ThioU46
U6ThioU51
UV
No cold pre-mRNA
U6
PRP2
chase product
crosslinks
L-IVS
pre-mRNA
1 2 3 4 5 6 7 8 9

Discussion

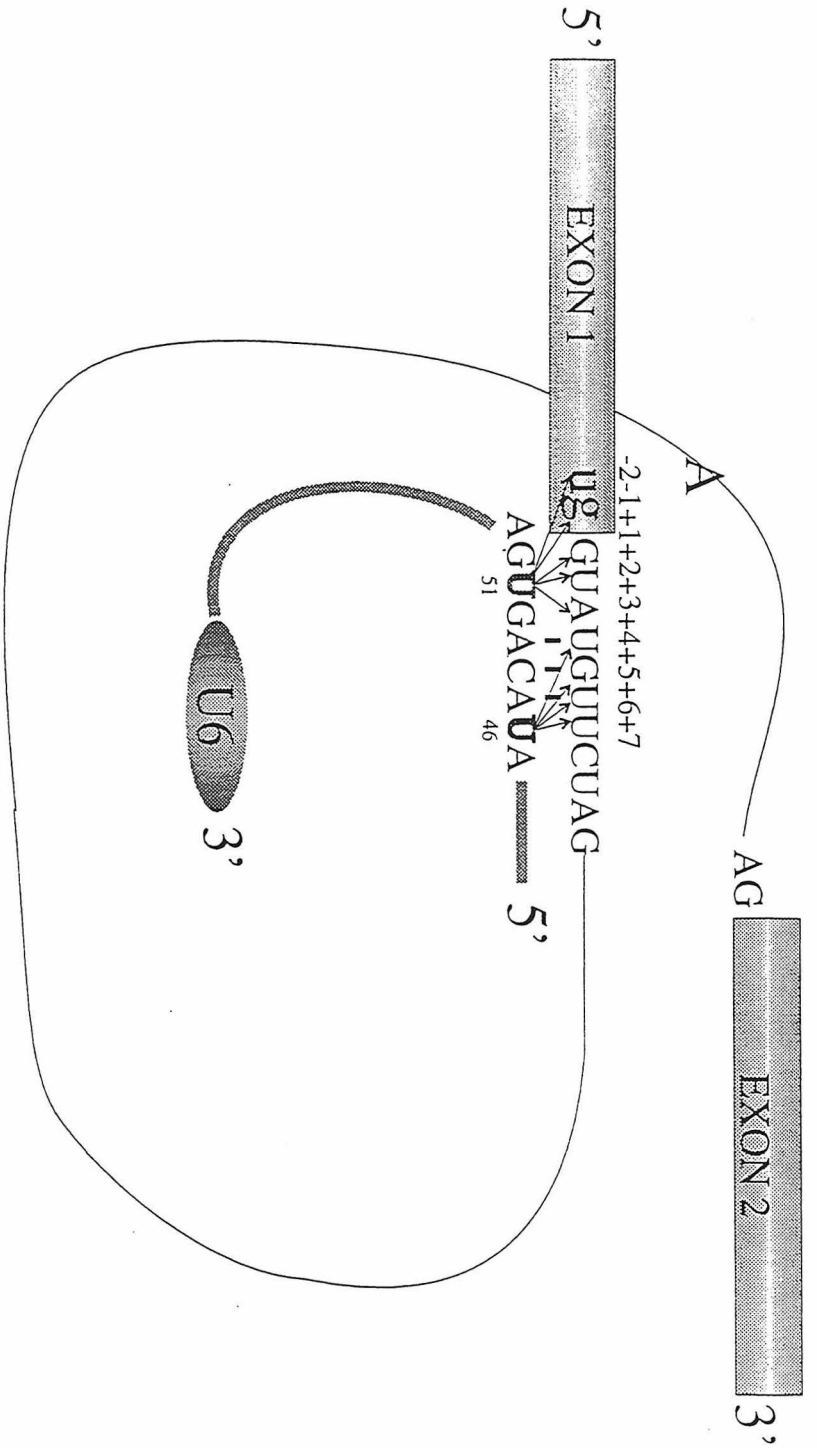
U6 RNA is an integral component of the dynamic network of RNA-RNA interactions that form the catalytic core of the spliceosome. These interactions take place after the U4/U6-U5 tri-snRNP joins the spliceosome and before the first catalytic step of splicing. We wish to more completely understand what RNA-RNA and RNA-protein interactions are required to assemble the catalytic core of the spliceosome. This paper demonstrates that a set of tools is available that can considerably expand our knowledge of what interactions occur during assembly and at what stage these interactions occur. We constructed four distinct U6 RNAs containing a single 4-thioU at positions 46, 51, 54 and 57 in the central domain of U6 RNA. All of these RNAs could restore activity to splicing extracts in which the endogenous U6 RNA had been inactivated. Upon photoactivation with UV, two of these RNAs, 4-thioU46 and 4-thioU51 formed specific crosslinks with the pre-mRNA or with the exon2-lariat intron reaction intermediate. The yield of these crosslinks was in some cases as high as 30% of the intron lariat intermediate formed during the splicing reaction. Such yields suggest that the crosslinks are not due to infrequent interactions. The high yields also make these experiments feasible. In an earlier study we observed UV dependent crosslinks between U6 RNA and the exon2-lariat intron intermediate (Sawa and Abelson, 1992) but the yields were at least 1000 fold lower making the mapping of the crosslinks extremely difficult.

From the results of reverse transcriptase mapping of the RNA species containing a crosslink between 4-thioU51 and the pre-mRNA, we infer that the crosslinks are to five positions in the pre-mRNA: to nucleotides -2, -1 (the last two bases of exon 1), and +1, +2,+3, (the first three bases of the intron). It seems unlikely that the stops are due to multiple reverse transcriptase stops in response to a single crosslink. The block to reverse transcriptase by the branched nucleotide results in one or two bands and more significantly, reverse transcriptase mapping of the 4-thioU51 crosslink to the exon2-intron lariat

intermediate revealed only two blocks. Most likely crosslink 1 (Fig. 3) is a mixture five different crosslinked species and each of these was equally effective as a template for primer-directed reverse transcription.

The crosslinks between 4-thioU51 and the intermediate may be directly compared to the results of Sontheimer and Steitz (1993). With 4-thioU incorporated uniquely at position two of the intron, they observed an efficient crosslink between the exon2-intron lariat intermediate and A45 in U6 RNA, the mammalian equivalent of yeast A51. Thus crosslink 3 in Figure 3 is the reverse of the Sontheimer and Steitz crosslink. Interestingly they only observed two crosslinked species, predominantly to A45 and less efficiently to A43. Sontheimer and Steitz did not observe a crosslink between pre-mRNA and U6 but they did observe a crosslink between U6 and the excised intron. The intron crosslinks were less specific; four separate crosslinks were observed. We could not observe a crosslink between U6 and the lariat intron product because U51 blocks splicing at the second step. An interpretation that could be drawn from the two studies is that the interaction between U6 and the intron is looser before the first step of splicing, very constrained after the first step, and again less constrained after the reaction is complete. Irradiation of extracts reactivated by addition of 4-thioU46 U6 resulted in four different crosslinks to intron positions +4,+5,+6 and +7 of the pre-mRNA. Though bands consistent with the mobilities of the 4-thioU51 crosslinks 2 and 3 were seen, the yields of these crosslinks were too low to obtain reverse transcriptase mapping data. The positions of the U46 crosslinks are consistent with the yeast genetic studies of Lesser and Guthrie (1993) and Kandels-Lewis and Seraphin (1993) which suggested a functional base-pairing interaction between the $A_{47}C_{48}A_{49}$ sequence in U6 and the $U_4G_5U_6$ sequence in the intron (Figure 1B). Taken together with the 4-thioU51 result it appears that there is a close interaction between U6 and the intron that extends between -2, the second last base of exon 1, and +7 of the intron and between A51 in U6 and at least A46 (Figure 10). If ten bases are also involved, the U6 interaction with the intron could span the distance between A51

Figure 10. Site-specific crosslinks of U6 to pre-mRNA near 5'splice site. Arrows indicate the sites of 4-thioU crosslinks from U6 to pre-mRNA. A 4-thioU at position 46 several nucleotides upstream of position 51 in U6 forms crosslinks to approximately the same number of nucleotides downstream in the intron of the pre-mRNA. This suggests an antiparallel configuration of the U6-5' splice site interaction required for base-pairing interaction. Genetically proven base-pairing between U6 and 5' splice site indicated by dashes (Lesser and Guthrie, 1993; Kandels-Lewis and Seraphin 1993) lie within the region defined by the crosslinks.


5'
EXON 1
-2-1+1+2+3+4+5+6+7
GUAUGUUCUAG
AGUGACAUA
51
46
5'
U6
3'
A
AG
EXON 2
3'

and A42. This interaction would rarely consist of canonical base-pairs. The only conserved base-pairs among yeast introns are between $U_4G_5U_6$ and $A_{47}C_{48}A_{49}$ and in mammalian introns the only conserved base-pair interactions are with U_4 and G_5.

In an earlier study (Sawa and Abelson, 1992) we observed high energy UV crosslinks between the exon2 lariat intron intermediate and U6 RNA. These interactions were mapped to residues G39, A41, A42, C43, A44 in U6 and nucleotides +4, +5, +6,+7,+8 in the intron. These crosslinks are out of register with the 4-thioU46 crosslinks observed here which are to the overlapping bases +4,+5 +6, and +7. However, the two results cannot be directly compared since the 4-thioU46 cross-link is to the pre-mRNA and the UV cross-links were to the intermediate. The chemical mechanism of UV-induced cross-linking is different from that of the low energy activated 4-thioU crosslink though both mechanisms favor base-stacking interactions (Favre, 1990). The UV induced crosslinks were extremely rare and may have occurred only in a rarely populated interaction state, for example during a structural rearrangement taking place between the first and second steps.

The crosslinks between 4-thioU51 and 4-thioU46 and the intron occur efficiently in prp2-1 heat inactivated extracts. These extracts are blocked in spliceosome assembly just prior to the first chemical step. So the U6-intron interactions occur before that step. It is likely that they first occur much earlier at the stage when the U4/U6-U5 tri-snRNP joins the pre-spliceosome. At that stage the U6 intron interactions probably replace the overlapping U1 base-pairs, destabilizing the U1 interaction with the spliceosome. A number of recent studies in the mammalian system show that the U1 interaction is not absolutely essential. A U2-U4/U6-U5 complex associates with a 5'-splice site oligonucleotide when the 5' end of U1 RNA is blocked (Konforti et al., 1993). It seems very likely that this interaction is with U6 RNA. Trans-splicing between a 5' splice site oligonucleotide and the 3' portion of an intron occurs when U1 RNA is inactivated and is stimulated by the addition of an excess of SR proteins (Chiara and Reed, 1995; Konforti and Konarska, 1995). In mammalian

extracts, U1 independent splicing can be observed if an excess of SR proteins is added (Crispino et al., 1993; Tarn and Steitz, 1994). All of these observations suggest that the pre-spliceosome is poised to enter into the intron U6 interactions and that these interactions are stabilized by SR proteins. However, it has not been proved that the close RNA-RNA interactions that allow the 4-thioU mediated crosslinks we have observed occur at an early stage of spliceosome assembly. Unfortunately we have very few mutants, if any, which block spliceosome assembly after the tri-snRNP has joined the spliceosome and before the U4 dissociation. (complex A2-1, Figure 1). Prp19 may mediate that step (Tarn et al., 1993) and one allele of Prp 4 may also block assembly at that step (J. Banroque, personal communication). As more genes required for yeast mRNA splicing are isolated it is likely that they will provide a route to a more precise and detailed dissection of the spliceosome assembly pathway.

The U6/pre-mRNA crosslink formed in prp2-1 heat inactivated mutant extracts can be chased into the U6/lariat intermediate species upon addition of Prp 2 protein, especially in the case of U6 4-thioU51. This is the first strong evidence that the U6-5'splice site interaction, suggested by genetic and biochemical experiments, (see introduction) is physically maintained during the first step of splicing. The comparable chase experiment, to demonstrate the functional relevance of the U6-5' splice site interactions, was not carried out in the mammalian 4-thioU crosslinking experiments (Sontheimer and Steitz 1993) or the psoralen crosslinking studies (Wassarman and Steitz 1992). However, functional crosslinks were demonstrated, via a chase experiment, between U5 and 4-thioU at position -1 in the 5' exon adjacent to the 5' splice site (Sontheimer and Steitz 1993). Interestingly, the physical covalent crosslinking of an invariant nucleotide in the the conserved ACAGAGA sequence does not interfere with the active site structure required for the first step of splicing because the crosslinked molecule is capable of going through the first step. Since the 5' splice site is clearly involved in the first chemical reaction, it seems likely that the geometry of the crosslink is compatible with the structure of the active site (in fact that

this geometry led to the efficient crosslink) rather than the alternative possibility that the crosslink does not interfere because it is some distance from the active site. There are clearly a number of crucial interactions occuring between snRNAs and the 5' splice site. We have shown that the U6-5' splice site interaction occurs prior to the Prp2 mediated step. In the mammalian system, the functional U5-5' splice site crosslink does not occur in splicing extracts blocked with EDTA (a treatment thought to arrest spliceosome assembly at the same stage as the prp2-1 block). This comparison hints at a precise pathway of interactions leading to the assembly of the active site. It will be very interesting to investigate the crosslinks between U5 RNA and the 5'splice site, as Newman et al. (1995) have recently done, using prp2-1 extracts.

The method by which we incorporated specific 4-thioU residues in U6 RNA is not a general one since it depended on the fortuitous coincidence of there being only three U residues in an important region of U6 RNA, none of them being essential. Furthermore the method is no longer required because Peter G. Stockley's laboratory has developed a synthesis of 4-thiouridine phosphoramidites allowing the chemical synthesis of oligonucleotides containing 4-thioU (Adams et al., 1994). We have entered into a collaboration with Stockley's laboratory to synthesize and use 4-thioU containing oligonucleotides for a more extensive survey of possible interactions between U6 RNA and RNA or proteins in the spliceosome. In a preliminary control for that collaboration, the B fragment containing a 4-thioU at position 51 (Fig. 2) was chemically synthesized by J. A. Murray and C. Adams in Stockley's laboratory. Unlike the oligonucleotide used in our experiments, the synthetic oligonucleotide had the wild type sequence except for 4-thioU51. This oligonucleotide was efficiently joined to fragments A and C in the ligation reaction. In fact we find that synthetic oligonucleotides are more efficiently ligated than transcription products, probably because they have no 3'-end heterogeneity. The ligated RNA was fully active in complementing U6 depleted extracts, and upon irradiation the same three crosslinked products were observed as seen in Figure 3. The way is now open

to incorporate 4-thioU or other crosslinking agents at specific sites not only in U6 RNA but in U2, U5 and in the substrate (Stade et al., 1989). By labeling the complementing RNA instead of the substrate, snRNA-snRNA interactions can be explored. An increasing number of mutant extracts will be available to delimit the stage at which a particular crosslink occurs. With these tools an ever more precise picture of spliceosome assembly should emerge.

In classic enzymology, explorations of the active site of enzymes were first done by chemically modifying the substrate because it was not possible, in general, to modify the enzyme. It is now possible, with difficulty, to introduce a variety of chemical substitutions to proteins (Noren et al., 1989). There is now substantial evidence that the active site of the spliceosome is composed of parts of four RNA molecules, the pre-mRNA substrate, U2, U6 and U5 snRNA. With the increasing sophistication and efficiency of RNA synthesis via phosphoramidite chemistry it will be possible to synthesize a wide variety chemical variants of these sequences and via reconstitution in vitro to assemble them into the active site. These variants can probe aspects of nucleotide bases, backbone phosphate groups, and ribose rings. In a recent study, for example, we have assembled 50 different U6 RNA molecules, entirely composed of synthetic oligonucleotides, each containing a deoxy nucleotide at a different position between residues 39 and 88 (Kim et al. submitted for publication). Using the complementation assay we were able to determine which 2'-hydroxyl groups in this important central region of U6 are essential for function. Only four positions were essential and one of these was in the important nucleoside, A51. A deoxynucleotide at position 51 blocks the first step of splicing. Interestingly all base changes in A51 allow the first step to occur and block the second step. With an increasing number of ribonucleotide phosphoramidites being synthesized and the ability to use deoxynucleotide substitutions at most sites in U6, we should be able to obtain a much more detailed picture of the chemical and structural features of the active site of this enzyme.

Material and Methods

Synthesis of 4-thiouridine triphosphate

For incorporation of 4-thioU at a single specific site by transcription with T7 RNA polymerase, 4-thioU triphosphate was synthesized from 4-thioU diphosphate (SIGMA) using diphosphate kinase and ATP (Stade et al., 1989) coupled with an ATP co-factor regenerating system based on phosphoenolpyruvate (PEP) as phosphoryl group donor activated by pyruvate kinase (Simon et al., 1990). Reaction mixtures (1ml) containing 10 mM 4-thio-UDP , 20 mM PEP (SIGMA), 0.1 mM ATP pH 8.0 (used as co-factor in catalytic amounts), 50 mM Tris-HCl pH 7.8, 2.5 mM $MgCl_2$, 0.1 mM EDTA, 5.0 mM dithiothreitol and 1.0 mM spermidine were incubated over night at room temperature in the dark (in brown Eppendorf tubes) with 50 units of nucleoside 5' diphosphate kinase (SIGMA) and pyruvate kinase (Boehringer Manneheim). The reaction was stopped by adding EDTA to a final concentration of 5 mM and SDS to a concentration of 0.1%, followed by incubation in 1mg/ml proteinase K (Boehringer Manneheim) for 15 min at 37°C. The original volume was split into two 0.5ml aliquots, sodium acetate added to 300 mM final concentration and SDS to 0.1% and the mixture was phenol extracted. The nucleotides were precipitated with 1 vol of acetone/ethanol over night at -20°C. TLC analysis on PEI plates using 0.75M sodium phosphate buffer at pH 3.4 showed the formation of a new spot, corresponding to 4-thioUTP. Using the ATP regeneration system approximately 90% of the 4-thio-UDP was converted to 4-thio-UTP, which is more efficient than the 50% yield obtained in the absence of the ATP regeneration system (Stade et al., 1989). This nucleotide mixture (containing mostly, 4-thio-UTP, small amounts of ADP and ATP, and PEP) was used in the T7 transcription reactions without further purification.

Synthesis of U6 RNA

The control B transcripts (see Fig. 2) containing the necessary double or quadruple mutations but no 4-thioU were synthesized in a 600 µl reaction volume containing 40 mM Tris-HCl (pH 8.0), 1 mM spermidine, 5 mM DTT, 8% polyethylene glycol (PEG), 4mM GMP, 1mM of each ribonucleoside triphosphate, 14 mM $MgCl_2$, 0.3 µM DNA template (hybridized with the T7 primer template by annealing at 90^0C), 80 µCi α-^{32}P GTP and 25 U/µl of T7 RNA polymerase (NEB). The 4-thioU-containing B fragments were made in 100 µl reaction volume containing the 4-thioUTP nucleotide mixture described above, 4 mM each of CTP, ATP, GTP, 16 mM GMP, 60 µCi α-^{32}P CTP, 8% PEG, 5 mM DTT, 1 mM spermidine, 14 mM $MgCl_2$, 0.3 µM T7 primer-hybridized DNA template, and 25 U/µl of T7 RNA polymerase. The B and C fragments must be monophosphorylated at the 5' end to be a suitable phosphoryl group donor for the ligation, and this was achieved by transcribing with a four-fold excess of guanosine monophosphate (GMP) over the other nucleotide triphosphates to "cap" the 5' end with GMP (Milligan and Uhlenbeck, 1989). The reaction mixtures were incubated for four hours at 37 ^{o}C. Reactions were stopped by the addition of EDTA to a final concentration of 15 mM, extracted with phenol/chloroform/isoamyl alcohol (50:49:1) (PCA), ethanol precipitated, and purified by electrophoresis in a 15% polyacrylamide (19:1 acrylamide to bis) 8 M urea gel. Runoff transcripts initiate at a unique site, but may be heterogeneous at the 3' terminus -- all bands were tested for ligation. Following autoradiography, the bands were cut out from the gel, crushed and eluted in an elution buffer [0.5 M sodium phosphate buffer, 50 mM Tris-HCl (pH 7.5), 1 mM EDTA, 0.1% SDS] by mixing on a gentle rocker at 4^{o}C for 6hrs. The gel fragments were then removed by centrifuging in paper-filter columns (ISOLAB) and the resulting filtrate was PCA extracted, precipitated in ethanol, resuspended in distilled water and stored at -20 ^{o}C. The yields from the transcription reactions were quite high -- in the hundreds of pmole to nanomole range.

The A and C fragments were made in larger volume -- 3ml. The same concentrations of reagents and recovery methods from the gel were used as for the transcription of the wild-type B fragment. However, A and C fragments were not radioactively labeled to prolong shelf life. Thus, they had to be cut out from the gel with the help of a UV-lamp to illuminate the bands containing the RNA. The RNA concentrations were determined by measuring optical density at 260nm.

Synthetic U6 snRNA were made by the ligation of the three fragments A, B and C in a total volume of 10 µl. Each of the three fragments was annealed to a DNA-bridging oligodeoxynucleotide complementary to positions 20 to 100 on the U6, and subsequently ligated using T4 DNA ligase. [Because T4 DNA ligase requires precisely base-paired junctions, transcripts with non-template-encoded nucleotides at the 3' end (Milligan and Uhlenbeck, 1989) are not ligation acceptors]. 80 pmoles of each of the three short fragments of RNA and 60 pmoles of the bridging deoxynucleotide dissolved in a minimum amount of solvent were heated to 90°C for 2 min and cooled to room temperature in a water bath over 10 min. Reaction volume was then brought up to 9 µl, and 1 µl of the USB T4 DNA ligase buffer [0.5 M Tris-HCl (pH 7.5), 10 mM ATP, 10 mM $MgCl_2$, and 1 mg/ml of acetylated bovine serum albumin (BSA)] and 1 µl of 10 U/µl (high concentration) T4 DNA ligase enzyme (USB-Amersham) were added. The reaction mixture was incubated at 37 °C for 3 hrs. At the end of the incubation, 10 µl of gel loading buffer (22 g urea, 194 µl H_2O, 4 µl 2% BPB and XC dyes, 2 µl 10x TBE heated and then finally dissolved with an additional 100 µl of water) was added to the reaction mix and the mixture was purified on a 10% polyacrylamide 8 M urea gel. The ligation efficiency was approximately 30%. After autoradiography, the band corresponding to full-length U6 snRNA was excised and eluted using the same conditions for the elution of B fragment described above, but for 12 hrs, after which the fragments were micro-filtered, PCA extracted, ethanol precipitated and resuspended in distilled water.

Splicing Reactions and Crosslinking

Yeast whole-cell extracts were prepared from either the protease-deficient EJ101 or the prp2-1 strain (in the latter case, protease inhibitors were added) according to Lin et al. (1985). Endogenous U6 RNA was destroyed by incubating the splicing extract for 30 min at 37°C with the deoxynucleotide d1 complementary to sequences (28-54) in U6 (Fabrizio and Abelson, 1990). For prp2-1 mutant extracts, the Prp2 protein itself was heat-inactivated during this incubation with d1, resulting in a simultaneous knock-out of both Prp2 and U6. Typical reactions contained 2 μl of splicing extract in 4 μl reaction mixtures containing 2.5 mM $MgCl_2$, 3% PEG-8000, 60 mM potassium phosphate (pH 7.0), 2 mM spermidine, and 0.3μM oligonucleotide in the presence of 2 mM ATP. No exogenous RNaseH was added because yeast splicing extract contains RNaseH activity. The wild-type splicing extract control was also incubated using the same conditions except in the absence of the deoxynucleotide. For in vitro functional assay, after treatment of the extract with oligodeoxynucleotide d1, 0.5 μl of in vitro-synthesized U6 RNA (transcribed entirely from a plasmid DNA construct or ligated from the three short fragments) was added and allowed to assemble for 5-10 min at 23 °C. Then 1 fmole of the ^{32}P-labeled actin precursor mRNA was added in a volume of 0.5 μl of the same buffer as above, except that the concentration of the reagents was 2x. In the case of crosslinking experiments using heat-inactivated prp2-1 mutant extracts functional Prp 2 protein was obtained from R. J. Lin. For the Prp2 protein chase experiment using U6/pre-mRNA crosslinks pre-formed in prp2-1 extracts, 150 fmoles of excess cold pre-mRNA transcript was used to block assembly of new spliceosomal complexes. All samples were incubated at 23 °C. After 5 minutes of splicing, the 5 μl reaction mixtures were transferred to individual wells on a micro-titer dish on ice and irradiated with a LASER light at 325nm (Liconix) at 9mW for 5 min. The reactions were terminated and deproteinized as described by Lin et al. (1985),

and loaded on a thin 8% polyacrylamide gel (29:1 acrylamide to bis) for analysis of the crosslinked/uncrosslinked RNA intermediates and products.

RNAse H and Primer extension

For analytical RNAse H digestions, RNA from the splicing and crosslinking reactions were ethanol precipitated twice to remove traces of 10% SDS from the proteinase K digest which could act as an inhibitor. The ^{32}-P labeled species crosslinked to U6 was identified as pre-mRNA, lariat intermediate or lariat intron by oligonucleotide directed RNAse H cleavage using oligonucleotides complementary to sequences in the 5' exon, intron and 3' exon. RNAse H digestion conditions were as used by Sawa and Shimura (1992). To identify the sites of crosslinks on the pre-mRNA or lariat intermediate, 40 fmol of the 5'-end labeled intron primer complementary to nucleotides +39 to +55 in the intron was annealed to 0.1 fmoles of crosslinked RNA in a total volume of 5 µl containing 1 µl of 5x hybridization buffer (1.5 M NaCl, 50 mM Tris-Cl, pH 7.5, 10 mM EDTA) at 65°C for 3 min and slowly cooled to room temperature over 1 hr. The annealing reactions were diluted with 25 µl of pre-warmed 1.25x reverse transcription cocktail (1.25 mM dNTPs (Boehringer Manneheim), 12.5 mM DTT, 12.5 mM Tris-HCl, pH 8.4 and 7.5 mM $MgCl_2$) and incubated with 1µl of 0.25 U/µl of avian myeloblastosis virus reverse transcriptase (Promega) for 30 min at 37 °C. The reactions were stopped with 0.5µl of 0.5M EDTA and incubated with 0.5µl of 2N NaOH at 60°C for 1hr to hydrolyze any radioactively labeled RNA template that could be misinterpreted as a primer extension stop. DNA from the extension reaction was then phenol extracted with phenol/chloroform/isoamyl (25:24:1) equilibrated at pH 8.0 to remove the reverse transcriptase. The aqueous layer was ethanol precipitated with 0.3M NaOAc. Dideoxy sequencing ladders were generated from plasmid encoding the actin pre-mRNA gene using

Sequenase (Amersham-USB). All reactions were fractionated and analyzed on an 8% polyacrylamide-7M urea denaturing gel.

Acknowledgments

We thank Sang-Ho Kim, Shueh-Lee Yuan, Dong-Ho Kim and Ren Jiang Lin (City of Hope) for their generous gift of Prp2 protein and advice, Christine Guthrie and Jon Staley for reading the manuscript critically, and Andrea Ghetti for Fig. 1A. This work was supported by NIH grant GM32637.

References

Adams CJ, Murray JB, Arnold JRP, Stockley PG. 1994. A convenient synthesis of S-cyanoethyl-protected 4-thiouridine and its incorporation into oligoribonucleotides. *Tetrahedron Letters 35:* 765-768.

Ares M Jr, Weiser B. 1995. Rearrangement of snRNA structure during assembly and function of the spliceosome. *Progress in Nucleic Acid Research and Molecular Biology 50* 131-159.

Brow DA, Vidaver RM. 1995. An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. *RNA 1*: 122-131.

Chiara MD, Reed R. 1995. A two-step mechanism for 5' and 3' splice-site pairing. *Nature 375*: 510-513.

Cortes JJ, Sontheimer EJ, Seiwert SD, Steitz JA. 1993. Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites in-vivo. *EMBO J 12:* 5181-5189.

Crispino JD, Blencowe BJ, Sharp PA. 1993. Complementation by SR proteins of pre-mRNA splicing reactions depleted of U1 snRNP. *Science 265*: 1866-1869.

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

Fabrizio P, McPheeters DS, Abelson J. 1989. In vitro assembly of yeast U6 snRNP: a functional assay. *Genes & Dev. 3*: 2137-2149.

Fabrizio P, Abelson J. 1990. Two domains of yeast U6 small nuclear RNA required for both steps of Nuclear Precursor Messenger RNA Splicing. *Science 250*: 404-409.

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

Favre A. 1990. 4-Thiouridine as an intrinsic photoaffinity probe of nucleic acid structure and interactions. In: Morrison H, ed. *Bioorganic photochemistry: Photochemistry and the nucleic acids* . New York: John Wiley & Sons. pp 379-425.

Favre A, Dubreuil YL, Fourrey, J-L. 1991. Photocoupling of the Intrinsic Photoaffinity Probe 4-Thiouridine to Nucleosides and Oligonucleotides. *New J Chem 15*: 593-599.

Fortner DM, Troy RG, Brow DA. 1994. A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. *Genes & Dev 8*: 221-233.

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

Kandels-Lewis S, Seraphin B. 1993. Role of U6 RNA in 5' splice site selection. *Science 262*: 2035-2038.

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

Konforti BB, Koziolkiewicz MJ, Konarska MM. 1993. Disruption of base pairing between the 5' splice site and the 5' end of U1 snRNA is required for spliceosome assembly. *Cell 75*: 863-873.

Konforti BB, Konarska MM. 1995. A short 5' splice site RNA oligonucleotide can participate in both steps of splicing in mammalian extracts. *RNA 1*: 815-827.

Legrain P, Chanfreau G. 1994. Pre-mRNA splicing: from intron recognition to catalysis. *Bull Inst Pasteur 92*: 153-179.

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

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

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

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

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

McPheeters DS, Fabrizio P, Abelson J. 1989. In vitro reconstitution of functional yeast U2 snRNPs. *Genes & Dev 3*: 2124-2136.

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

Moore MJ, Sharp PA. 1992. Site-specifically modified pre-mRNA: Substitutions of the 2'-hydroxyl groups at the splice sites. *Science 256*: 992-997.

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

Newman A, Norman C. 1992. U5 snRNA interacts with exon sequences at 5' and 3' splice sites. *Cell 65*:115-123.

Newman A. 1994. Small nuclear RNAs and pre-mRNA splicing. *Current Opinion in Cell Biology 6*: 360-367.

Newman AJ, Teigelkamp S, Beggs JD. 1995. snRNA interactions at 5' and 3' splice sites monitored by photoactivated crosslinking in yeast spliceosomes. *RNA 1*: 968-980.

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

Noren CJ, Anthonycahill SJ, Griffith MC, Schultz PG. 1989. A general method for site-specific incorporation of unnatural amino-acids into proteins. *Science 244* : pp182-188.

Sawa H, Shimura Y. 1992. Association of U6 snRNA with the 5' splice site region of pre-mRNA in the spliceosome. *Genes & Dev 6*: 244-254.

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

Simon E, Grabowski S, Whitesides GM. Convenient syntheses of cytidine 5'-triphosphate, guanosine 5'-triphosphate, and uridine 5'-triphosphate and their use in the preparation of UDP-glucose, UDP-glucuronic acid and GDP-mannose. *J Org Chem. 55*: 1834-1841.

Sontheimer EJ. 1994. Site-specific RNA crosslinking with 4-thiouridine. *Molecular Biology Reports 20*: 35-44.

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

Stade K, Rinke-Appel J, Brimacombe R. 1989. Site-directed cross-linking of mRNA analogues to the *Escherichia coli* ribosome; identification of 30S ribosomal components that can be cross-linked to the mRNA at various points 5' with respect to the decoding site. *Nucleic Acids Res 17*: 9889-9908.

Tarn TY, Lee KR, Cheng SC. 1993. Yeast precursor mRNA processing protein Prp 19 associates with the spliceosome concomitant with or just after dissociation of U4 small nuclear RNA. *Proc Natl Acad Sci USA 90*:10821-10825.

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

Umen JG, Guthrie C. 1995. The second catalytic step of pre-mRNA splicing. *RNA 1*: 869-885.

Wassarman DA, Steitz JA. 1992. Interactions of small nuclear RNAs with precursor messenger RNA during in vitro splicing. *Science 257:* 1918-1925.

Wu J, Manley J. 1991. Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. *Nature* (Lond.) *352*: 818-821.

Wyatt JR, Sontheimer EJ, Steitz JA. 1992. Site-specific cross-linking of mammalian U5 snRNP to the 5' splice site before the first step of pre-mRNA splicing. *Genes & Dev 6*: 2542-2553.

Yan D, Ares M. 1996. Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits. *Mol Cell Biol 16*: 818-828.

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

Chapter IV: Site-specific 4-thioU crosslinks of U6 snRNA to U4, U2 and the pre-mRNA

(Chang Hee Kim*, James B. Murray+, Chris J. Adams+, Peter G. Stockley+, John Abelson*.

Division of Biology, California Institute of Technology, Pasadena, California 91125*.

University of Leeds, Department of Genetics, Leeds LS29JT, W Yorkshire, England+.

To be submitted to Journal of Biological Chemistry, 1998)

Abstract

The active site of spliceosomes is assembled de novo around the pre-mRNA substrate, and is thought to include U6, U2 and U5 snRNAs. We have done an extensive search for crosslinks from 4-thioU site-specifically incorporated into U6 snRNA to other snRNAs or to the pre-mRNA substrate in order to probe the structure of the active site. We were able to look for crosslinks between U6 snRNA and other snRNAs because we have an in vitro system for functionally reconstituting U6 snRNPs from synthetic U6 RNA. We found ATP-dependent crosslinks to U4 snRNA from U6 RNA substituted with 4-thioU at nucleotides 51 and 80. We also found splicing-dependent crosslinks to U2 snRNA from U6 substituted with 4-thioU at positions 62 and 64. Interestingly, the 4-thioU51 and 4-thioU80 in U6 which crosslink to U4 snRNA also crosslink to the pre-mRNA in a splicing-dependent manner, supporting the idea that the interaction of U6 snRNA in the spliceosome is dynamic in nature, with the U4/U6 interaction occurring before the pre-mRNA / U6 interaction. We had previously reported that 4-thioU51 in U6 snRNA crosslinks to the pre-mRNA at the 5' splice site. In this study, we have found a novel splicing-dependent crosslink from 4-thioU80 in U6 RNA to A210 in the actin pre-mRNA

which resides in intronic sequences upstream of the branchpoint that are required for its splicing in vitro. These intronic sequences are thought to base-pair with complementary sequences just downstream of the 5' splices site, thereby facilitating the juxtaposition of the branchpoint adenosine with the 5' splice site, which is required for the first step of splicing. This is the first time that the conserved bulged U80 in the 3' U6 intramolecular stem-loop has been shown to interact with a component of the spliceosome.

Introduction

A complex network of RNA-RNA interactions is thought to play a key role in splice site recognition and catalysis of pre-mRNA splicing (reviewed in (Ares and Weiser, 1995; Legrain and Chanfreau, 1994; Madhani and Guthrie, 1994a; Newman, 1994; Nilsen, 1994; Nilsen, 1998; Staley and Guthrie, 1998). The speculation that pre-mRNA splicing may be catalyzed by the snRNA components is based mainly on similarities between pre-mRNA splicing and group I or group II intron self-splicing which can be catalyzed in the absence of any proteins (Cech, 1985; Sharp, 1985). Both pre-mRNA splicing and group II intron splicing proceed by two phosphodiester bond transesterifications, involving 1) the nucleophilic attack by the 2'-hydroxyl of the branchpoint A on the phosphate at the 5' splice site, resulting in the formation of a lariat intermediate with a 2'-5' linkage and a free, linear 5' exon, and 2) the nucleophilic attack by the 3' hydroxyl of the free 5' exon on the phosphate at the 3' splice site, resulting in the ligation of the two exons and release of the intron as a lariat by-product . Group I splicing and pre-mRNA splicing both have the same "in-line" SN2 nucleophilic displacement for both transesterification reactions, and the same stereochemical requirement for the Sp diastereomer for the second step. (Maschloff and Padgett, 1992; McSwiggen and Cech, 1989; Moore et al., 1993; Moore and Sharp, 1993; Rajagopal et al., 1989). Pre-mRNA splicing appears to utilize two different active sites

during the two steps of splicing, while group I intron splicing uses one active site and it is unclear whether group II introns splicing uses one or two active sites (Chanfreau and Jacquier, 1994; Moore et al., 1993; Moore and Sharp, 1992; Moore and Sharp, 1993; Sontheimer et al., 1997). Pre-mRNA splicing is thought have an active site similar to group II intron splicing for the first step, and to group I splicing for the second step(Moore et al., 1993). Furthermore, several of the snRNA interactions in the spliceosome appear to have counterparts in the group II intron (Cavalier-Smith, 1991; Moore et al., 1993; Sharp, 1991; Weiner, 1993; Wise, 1993). Thus, pre-mRNA splicing may be dependent on catalysis by RNA, with conserved snRNAs assembling the active site of the spliceosome through a network of RNA-RNA interactions.

Pre-mRNA splicing is executed via a number of assembly steps and chemical steps distinguished by the RNA contents of biochemically separable spliceosome complexes. At the molecular level, these assembly/chemical steps correspond to changes in RNA-RNA interactions during splice site recognition and catalysis. First, U1 snRNP interacts with the pre-mRNA to form the "commitment complex" in which the 5'-end of U1 snRNA base-pairs with the 5' splice site (Reich et al., 1992; Siliciano and Guthrie, 1988; Zhuang and Weiner, 1986). This is the only complex that forms in the absence of ATP. In the next step, U2 snRNP joins this complex to form the "pre-spliceosome" in an ATP-dependent manner. In this complex, the 5' end of U2 snRNA base-pairs with the conserved branchpoint sequence (Parker et al., 1987; Wu and Manley, 1989; Zhuang and Weiner, 1989). U4 and U6 are found in a single snRNP containing U4/U6 duplex held together by 24 base-pairs in two helices (Madhani et al., 1990). The U4/U6 snRNP is joined with U5 snRNP in a triple snRNP which enters the spliceosome in the next step. In this step, a number of RNA-RNA re-arrangements take place: 1) The 5'-end of the highly conserved ACAGAGA sequence in U6 interacts with the intronic sequences in the 5' splice site by base-pairing and adjacent U6 nucleotides can form crosslinks to the 5' splice site region (29-33). The base-paired interaction is mutually exclusive with that between U1 and the 5'

splices site. In fact, the U6/5' splice site interaction replaces the U1/5' splice site interaction and destabilizes(Chanfreau and Jacquier, 1994) U1 from the spliceosome. 2) Prior to the first step of splicing, three nucleotide bases in a conserved loop of U5 snRNA interact with last three bases of exon 1 (Cortes et al., 1993; Newman and Norman, 1992; Sontheimer and Steitz, 1993). However, this interaction and the loop in U5 is not required for the first step of splicing to proceed (O'Keefe et al., 1996). 3) The 3' end of U6 base-pairs with the 5'-end of U2 to form helix II which is important in mammalian cells(Datta and Weiner, 1991; Wu and Manley, 1991), but perhaps not in yeast (Fabrizio et al., 1989; Madhani et al., 1990). In the next assembly step, U4 dissociates from the spliceosome(Yean and Lin, 1991). The U4/U6 helices are disrupted and replaced by the formation of U2/U6 helix and an intramolecular stem-loop at the 3' -end of U6 (Brow and Vidaver, 1995; Fortner et al., 1994; Madhani and Guthrie, 1992). At this point, the active site of the spliceosome is ready for the first chemical step of splicing. The interactions between U6 snRNA and the 5' splice site, U2 snRNA and the branchpoint, and U6 and U2 with each other may facilitate the juxtaposition of the 5' splice site to the branchpoint A for catalysis of the first step of splicing. Much less is known about the RNA-RNA interactions that may mediate the second chemical step. Presumably, the conserved loop of U5 aligns the free 5' exon with the 3' exon on the lariat intermediate to assist in the exon-ligation step (Newman and Norman, 1992). Reconstitution of U5 snRNPs in vitro indicated that interactions between the U5 loop and 5' exon are critical for the second step of splicing (O'Keefe et al., 1996). In group II introns, the counterpart to the exon-binding loop of U5 snRNA -- the ID 3 stem-loop -- is also dispensable for the first step of splicing, but required for the second step (Newman, 1997). Intriguingly, the conserved stem-loop of human U5 snRNA can substitute for the group II intron ID3 stem-loop to mediate the second step of group II intron splicing in a trans-activation assay (Hetzer et al., 1997).

In summary, the RNA-based catalytic core of the spliceosome may be composed of U6 and U2 for the first chemical step, and U5 for the second step. U1 and U4 may be

excluded since they dissociate prior to the chemical steps. U6 RNA is a particularly good candidate for a catalytic RNA because: 1) it is the mostly highly conserved snRNA -- U6 RNA from rat, yeast, and plant are > 70% homologous (Brow and Guthrie, 1988); 2) extensive mutational analysis in a variety of systems has identified nucleotides in U6 that are essential for the first and/or second steps of splicing (Fabrizio and Abelson, 1990; Madhani and Guthrie, 1992; Vankan et al., 1990), reviewed in (Madhani and Guthrie, 1994a) ; 3) there are pre-mRNA type-introns in the middle of the most highly conserved region (ACAGAGA) in the U6 RNAs of fission yeast *Schizosaccharomyces pombe* (Tani and Ohshima, 1989); 4) in nematodes, alterations upstream of the ACAGAGA sequence in U6 causes an aberrant splicing reaction in which U6 itself is attacked by the branchpoint A of the pre-mRNA (Yu et al., 1993).

In order to elucidate the functional role of U6 RNA and to probe the U6 RNA-RNA interactions at the active site of the spliceosome, we have performed an extensive search for site-specific 4-thioU crosslinks from U6 to other snRNAs or to the pre-mRNA substrate. This was made possible by two technical capabilities. 1) We utilized our in vitro reconstitution system in which the endogenous U6 can be depleted and replaced by a synthetic U6 RNA (Fabrizio et al., 1989; Kim and Abelson, 1996), containing 4-thioU in this report. 2) We were able to collaborate with the Stockley lab in University of Leeds, Leeds, U.K. who originated the synthesis of 4-thioU phosphoramidite(Adams et al., 1994). We obtained short RNA oligonucleotides containing 4-thioU incorporated at a single site using phosphoramidite chemistry. These were subsequently ligated into full length synthetic U6. Prior to this technical breakthrough, we were restricted to testing only 4-thioU positions in U6 transcripts (Kim and Abelson, 1996).

Materials and Methods

Synthesis of U6 RNA

The U6 RNAs were synthesized from a ligation of the following four or three pieces: A1-38, B 39- 59, C60-79, D80-112; or A1-38, B39-79, D80-112, depending on which RNA had the 4-thioU modification. For the 4-thioU modifications in nucleotides 60-79, the four-piece ligation scheme was used, and for modifications in nucleotides 80. 112, the three-piece scheme was used. Modifications in the B fragment were introduced with the three-piece scheme since the ligation yields obtained with this method were higher. The synthetic RNA oligonucleotides with site-specifically incorporated 4-thiouridines (using phosphoramidite chemistry) were obtained in collaboration with the Stockley group in the University of Leeds. The rest of the RNA oligonucleotides were also chemically synthesized at the RNA synthesis facility at the Beckman Institute at Caltech or purchased from Oligos etc. The longer pieces of RNA were gel-purified on 10% polyacrylamide/7M urea gels to remove the shorter abort products from the chemical synthesis. In the four-piece ligation scheme, the 5'-ends of fragments B, C and D had to be mono-phosphorylated to act as donor groups in the ligation reactions. In the three-piece scheme, fragments B and D had to mono-phosphorylated at the 5'-ends. Two different types of labeling reactions were done depending on the level of specific activity desired. For high specific activity, the 5'-end labeling was done using 7000 Ci/mmole [γ-P32] ATP (ICN), in 20 μl reactions containing 80 pmoles of each of the RNA substrates, 100 pmoles of [γ-P32] ATP, 1x T4 polynucleotide kinase buffer [70 mM Tris-HCl pH. 7.6, 10 mM $MgCl_2$, 5 mM dithiothreitol (DTT)] (New England Biolabs, NEB) and 1 μl of 10 U/μl T4 polynucleotide kinase (NEB). For the low specific activity labeling, 20-fold less radioactive [γ-P32] ATP was used, and cold ATP was supplemented to 100 pmoles to ensure that the 5'-end of the phosphoryl donor oligonucleotides in the ligation reaction

were fully phosphorylated. The reaction mixture was incubated at 37°C for 30 mins, brought up to 100 µl with distilled and deionized water, P/C/A extracted, chloroform extracted, ethanol precipitated with 0.3 M NaOAc, and resuspended in water prior to use in the ligation reaction.

The kinased RNA oligonucleotides and the 5'-end A1..38 fragment were used as substrates for DNA template-assisted ligation using T4 DNA ligase (Moore and Sharp, 1992). Eighty picomoles of each of the three or four pieces of RNA was annealed to 60 picomoles of bridging DNA oligonucleotide complementary to positions 20-100 of U6 RNA by heating to 95°C and cooling slowly to room temperature over 5 mins. Reaction volume was then brought up to 8 µl by adding water, to which, 1 µl of T4 DNA 10x ligase buffer (0.5 M Tris-HCl, pH 7.5, 10 mM ATP, 10 mM $MgCl_2$, and 1 mg/ml of acetylated bovine serum albumin (BSA)) [USB-Amersham] and 1 µl of 10 U/µl (high concentration) T4 DNA ligase (USB-Amersham) were added. The reaction mixture was incubated at 37°C for 3 h. At the end of the incubation, 10 µl of denaturing gel loading buffer were added to the reaction mixture, and the mixture was purified on a 10% polyacrylamide(19:1 acrylamide to bisacrylamide)/ 7 M urea gel. The ligation efficiency was usually close to 80%. The gel bands corresponding to the full-length U6 RNA were excised, and the RNA was eluted from the crushed gel slice eluted by soaking in 500 µl of elution buffer (0.3M sodium phosphate buffer, 50 mM Tris-Cl (pH 7.5), 1 mM EDTA, 0.1%SDS). After gentle rocking at 4°C for 4 hrs, the elution mixture was spun down, purified by DEAE ion-exchange using Elu-r-tip (Schleicher & Schuell), ethanol precipitated and resuspended in distilled water.

Splicing Reactions and Crosslinking

Yeast whole-cell extracts were prepared from the protease-deficient EJ101 strain according to Lin et al. (Lin et al., 1985). Endogenous U6 RNA was destroyed via

RNAseH digestion by incubating the splicing extract for 30 min at 34°C with the deoxynucleotide d1 complementary to sequences (28-54) in U6 (Fabrizio and Abelson, 1990; Fabrizio et al., 1989). Typical reactions contained 2 µl of splicing extract in 4 µl reaction mixtures containing 2.5 mM $MgCl_2$, 3% PEG-8000, 60 mM potassium phosphate (pH 7.0), 2 mM spermidine, and 0.3-0.4 µM oligonucleotide in the presence of 2 mM ATP. No exogenous RNaseH was added because yeast splicing extract contains RNaseH activity. For the -ATP control experiments, ATP in the extract was depleted following d1 oligonucleotide treatment by incubation with 1 µl of 20 mM glucose for 5 min. at room temperature. The wild-type splicing extract control was also incubated using the same conditions except in the absence of the deoxynucleotide. To reconstitute extracts with in vitro synthesized U6 RNA (transcribed entirely from a plasmid DNA construct or ligated from chemically synthesized RNA oligonucleotides), 0.5 µl of 100 fmole/µl U6 RNA was added and allowed to assemble for 5 min at 23 °C. Then 1 fmole of the ^{32}P-labeled actin precursor mRNA was added in a volume of 0.5 µl of the same buffer as above, except that the concentration of the reagents was 2x. In experiments where unlabeled pre-mRNA was used (in order to look for RNA crosslinks to P-^{32} labeled U6), 1 fmole of unlabeled actin precursor pre-mRNA was added instead. All samples were incubated at 23 °C for 10 min, then the 5 µl reaction mixtures were transferred to individual wells on a micro-titer dish on ice and irradiated with a 365 nm UV lamp (Model B-100, UVP, Inc.) for 20 min. The source output at 325 nm was 5 mW. A petri dish was used as a low UV cut-off filter (<290 nm). The reactions were terminated and deproteinized as described previously (Lin et al., 1985), and loaded onto a 40 cm long, 0.4mm thick 5% polyacrylamide gel (19:1 acrylamide to bis) and run overnight for analysis of the crosslinked/uncrosslinked RNA intermediates and products.

RNAse H and Primer extension

For analytical RNAse H digestions, RNA from the splicing and crosslinking reactions were P/C/A extracted and ethanol precipitated twice to remove traces of SDS from the proteinase K digestion which could act as an inhibitor. The total RNA resulting from the splicing and crosslinking reactions was subjected to RNAseH digestion. The RNA crosslinks to U6 RNA was identified by oligonucleotide directed RNAse H cleavage using DNA oligonucleotides complementary to sequences in U1, U2, U4, U5, the 5' exon and the intron. Each cDNA oligonucleotide (1 µl of 0.2 µg/µl) was annealed to the recovered total RNA (resupended in 7 µl of distilled water) in a total volume of 20 µl with 12 µl of splicing buffer D (20 mM Hepes pH 7.5, 0.2 mM EDTA, 50 mM KCl, 0.5 mM DTT, 20% glycerol) by heating to 75^{o}C for 3 mins in a water bath, and slowly cooling to room temperature over 10 mins. After the annealing step, RNAseH digestions were carried out by adding 1 µl of 6 µg/µl tRNA, 2 µl of 25 mM $MgCl_2$ and 1 µl of 1 Unit/µl RNAseH (Takara) and incubating for 10 mins at 34 ^{o}C. Control reactions were incubated without any oligonucleotide and subjected to the same annealing procedure and incubation with RNAseH. After incubation, the RNAs were ethanol precipitated overnight, and separated on an 8% polyacrylamide (19:1 bis to acryl)/ 7 M urea sequencing gel to analyze the RNAse H digestion products.

For primer extension mapping experiments, the RNA crosslink was excised from the gel and electro-eluted in dialysis bags (12-14,000 molecular weight cut-off, Spectra/Pore) containing 1 ml of 1x TBE buffer. Electro-elution was carried out for 2 hours in 1x TBE buffer in a horizontal gel tank, after which polarity was reversed for 5 min to remove RNA sticking to the dialysis bag. The electro-eluted RNA was P/C/A and chloroform extracted, then ethanol precipitated overnight. Recovery of RNA via electro-elution was usually about 80%. To map the sites of crosslinking on the pre-mRNA, we used a primer, **3'F** (5' CGAGGAGCGTCGTCACC 3'), complementary to sequences in the 3'-exon of the

actin pre-mRNA. The 3' -F primer was labeled to high specific activity in a 10 µl volume including 6 µl of 7000Ci/mmol [P^{32}-γ] ATP (ICN) and 1 µl of 10 U/µl of T4 polynucleotide kinase (NEB) in 1x T4 polynucleotide kinase buffer (NEB) for 30 mins at 37°C. The unincorporated [P^{32}-γ] phosphates were removed by separating the reaction mixture on a short (7cm) 20% polyacrylamide (19:1 acryl: bis-acryl)/ 7 M urea denaturing gel. The excised band containing only the 5'-end labeled primer was crushed and eluted in elution buffer (same as above) for 15 mins at room temperature. The gel fragments were briefly spun down in a microcentrifuge, and the eluate was loaded onto a reverse-phase Sep-Pak C_{18} column cartridge (Fisher) for purification. The column is pre-washed with 5 ml of acetonitrile, and equilibrated with 10 ml of 100 mM TEAB(triethylammonium bicarbonate) buffer at pH 7.6 (acidified using carbon dioxide gas). The DNA eluate containing the primer is loaded as is, in a volume of 500 µl. The load is washed with 10 ml of water and eluted out using 2 ml of 50% acetonitrile which is evaporated to dryness overnight in a speedvac. The resulting powder is then resuspended in 20 µl of double-distilled water prior to use in primer extension reactions.

Forty fmoles of the gel-purified 5'-end labeled primer were then annealed to 0.1 fmoles of crosslinked RNA in a total volume of 5 µl containing 1 µl of 5x hybridization buffer (1.5 M NaCl, 50 mM Tris-HCl (pH 7.5), 10 mM EDTA) at 65°C for 3 min and slowly cooled to room temperature over 1 hr. The annealing reactions were diluted with 25 µl of pre-warmed 1.25x reverse transcription cocktail (1.25 mM dNTPs (Boehringer Manneheim), 12.5 mM DTT, 12.5 mM Tris-HCl, pH 8.4 and 7.5 mM $MgCl_2$) and incubated with 1 µl of 0.25 U/µl of avian myeloblastosis virus reverse transcriptase (Promega) for 30 min at 37 °C. The reactions were stopped with 0.5 µl of 0.5M EDTA and incubated with 1 µl of 2N NaOH at 60°C for 1 hr to hydrolyze radio-labeled RNA crosslink that could be misinterpreted as a primer extension stop. DNA from the extension reaction was then extracted with phenol/chloroform/isoamyl (25:24:1) equilibrated at pH 8.0 to remove the reverse transcriptase. The aqueous layer was ethanol precipitated

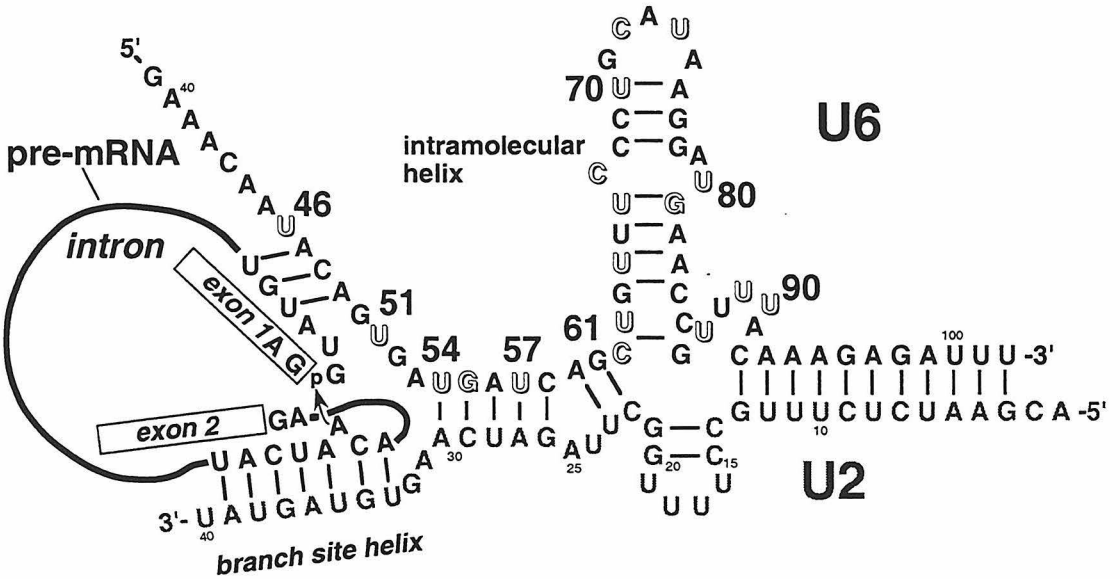
overnight. Dideoxy sequencing ladders were generated from plasmid encoding the actin pre-mRNA gene using Sequenase (Amersham-USB). All reactions were fractionated and analyzed on an 8% polyacrylamide-7M urea denaturing sequencing gel.

Results

Site specific incorporation of 4-thiouridine.

The positions of all the 4-thioU crosslinks tested in U6 RNA are shown in Fig. 1 The U6 RNA conformation within the fully assembled spliceosome is depicted in this figure. Prior to this conformation, U6 is found extensively base-paired with U4 in a U4/U6 snRNP or a U4/U6.U5 triple snRNP (see Fig. 7a). These nucleotides lie in or very near the highly conserved central domain of U6 (nucleotides 47-85 in yeast) (Brow and Guthrie, 1988; Roiha et al., 1989; Schumyatsky and Reddy, 1993). The only 4-thioU nucleotides which were studied outside the conserved domain were at positions U87, 89 and 90. Crosslinks from 4-thioU46, 51, 54 and 57 in U6 RNA to P^{32}-labeled pre-mRNA substrate were described previously (Kim and Abelson, 1996). However, 4-thioU crosslinks from these positions in U6 RNA to other spliceosomal RNAs were not tested mainly because we could not label the U6 RNA to very high specific activity. Moreover, only 4 positions were tested because 4-thioU could only be introduced by transcription of a short central fragment of U6 RNA using 4-thioUTP in place of UTP from a DNA template coding for a single uridine (Kim and Abelson, 1996). The wild-type sequence of this short fragment contained three U's, and because we wanted to introduce only one 4-thioU at a time, the other two U's had to be mutated. The entire length of the U6 RNA molecule was synthesized by ligation of the flanking sequences of U6 to the transcript containing 4-thioU.

Figure. 1 Positions of all 4-thioU crosslinks tested in yeast U6 snRNA. The conformation of U6 RNA in the fully assembled spliceosome is depicted. U6 RNA can also be found in another conformation early in spliceosome assembly, extensively base-paired to U4 (Fig. 7a). A single 4-thiouridine was introduced at each of the shadowed positions by ligating three or four short chemically synthesized RNA oligonucleotides, one of which contained a single 4-thioU introduced by phosphoramidite chemistry (see Materials and Methods for details). The wild-type sequences is A for 51, C for 61, and 66, 67 and 72. However, for the purpose of this figure, they are indicated as 4-thioU.


pre-mRNA
intron
exon 1
A G
exon 2
intramolecular
helix
U6
U2
branch site helix
5'
3'-
-3'
-5'

The chemical synthesis of the 4-thiouridine phosphoramidite was a technical breakthrough that enabled the incorporation of a single 4-thiouridine at a selected site in a short RNA oligonucleotide (up to 40 nucleotides long) (Adams et al., 1994). Through our collaboration with Peter Stockley's group in University of Leeds, we were able to obtain short synthetic RNA oligonucleotide pieces of U6 RNA that had single 4-thioU substituents at selected sites. Large, biologically interesting RNA molecules with a single 4-thiouridine can then be synthesized by ligation of the component pieces, one of which has a single 4-thioU substituent. Preparing U6 RNA entirely from synthetic pieces is feasible because U6 is relatively short (the full length wild type is only 112 nucleotides in length). Thus, we were able to substitute virtually any nucleotide in U6 RNA with a 4-thioU. In the present study, the entire length of U6 RNA was prepared via ligation of three or four pieces of synthetic RNA oligonucleotides, one of which contained a single 4-thiouridine incorporated by phosphoramidite chemistry (refer to Materials and Methods). In contrast to our previous studies (Kim and Abelson, 1996), we were able to prepare U6 RNA's with much higher specific activity by incorporating ^{32}P at the ligation junctions, and this substantially improved the lower limit of detection.

Introducing 4-thioU at some positions in U6 resulted in a bbase mutation: A51, C61, A62, C66, C67 and C72. We confirmed that these mutations did not result in loss of U6 RNA function (data not shown). We also found that a substitution of a wild-type U by 4-thioU was tolerated at any of the positions indicated in bold in Fig. 1 were tolerated. This may be because 4-thioU is analogous to U in its base-pairing properties, even though the thio-carbonyl bond is approximately 20% longer than the carbonyl bond length (Favre, 1990).

We were particularly interested in testing 4-thioU crosslinks from nucleotides in the conserved intramolecular stem-loop at the 3'-end of U6 (see Fig. 1). These include the nucleotides 61, 62, 64, 66, 67,70, 72, 74, 80, 81. Nucleotide 61 is also the last nucleotide in the "catalytic AGC triad" found in the 3' stem-loop of U6 RNA in pre-mRNA splicing,

and also in the analogous domain V helix in Group II splicing. However, the functional role of this conserved 3' stem-loop in pre-mRNA splicing is not known. The nucleotide at position 66 (C in wild-type) is a strong DMS protection site in spliceosome glycerol gradient fractions purified just prior to the catalytic fraction (S. L. Yean and R. J. Lin, personal communication). It is thought that this DMS protection may be a result of interaction between this nucleotide and a protein, or an essential RNA in the spliceosome. We have previously found 2'-deoxy substitutions that result in inhibition of splicing activity at positions 70 and 72 in the loop of the U6 3' stem loop (Kim et al., 1997). Uridine 80 in U6 RNA is a very highly conserved bulged nucleotide, and a phosphorothioate substitution 5' of this nucleotide or its analog results in loss of splicing in yeast, nematodes and in Group II splicing (Chanfreau and Jacquier, 1994; Fabrizio and Abelson, 1992; Yu et al., 1995). However, it is not yet known whether this conserved nucleotide interacts with some other essential component of the spliceosome. A 6-thioG was incorporated at G81 because of its proximity to the conserved U80 nucleotide. 6-ThioG was also placed at position 55 near the U2-U6 helix 1a. 6-ThioG is thought to crosslink to RNAs or proteins, but the mechanism of the crosslinking reaction is not as well known as that for 4-thioU (Adams et al., 1995). 4-ThioU's were also introduced to nucleotides 87, 89 and 90. These lie just outside of the conserved central domain of U6 RNA. The Peebles group had found crosslinks with 4-thioU at the analogous positions in domain V helix of Group II splicing introns (Podar et al., 1998). We wanted to investigate any similarities or differences between these two analogous systems.

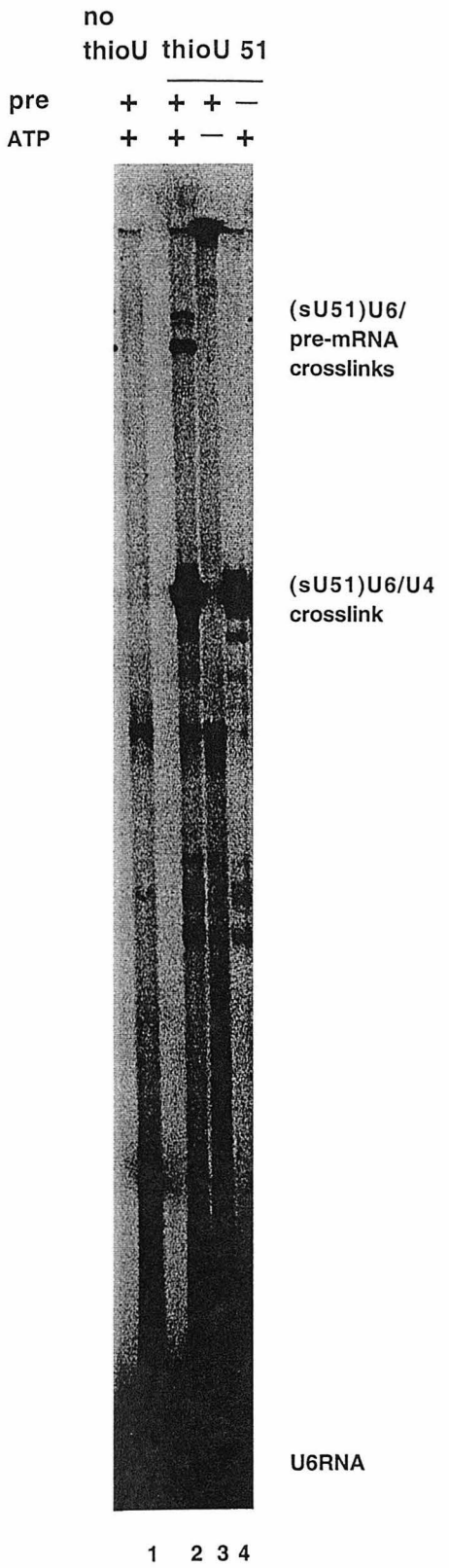
4-ThioU substituted U6 RNA crosslinks with spliceosomal RNAs

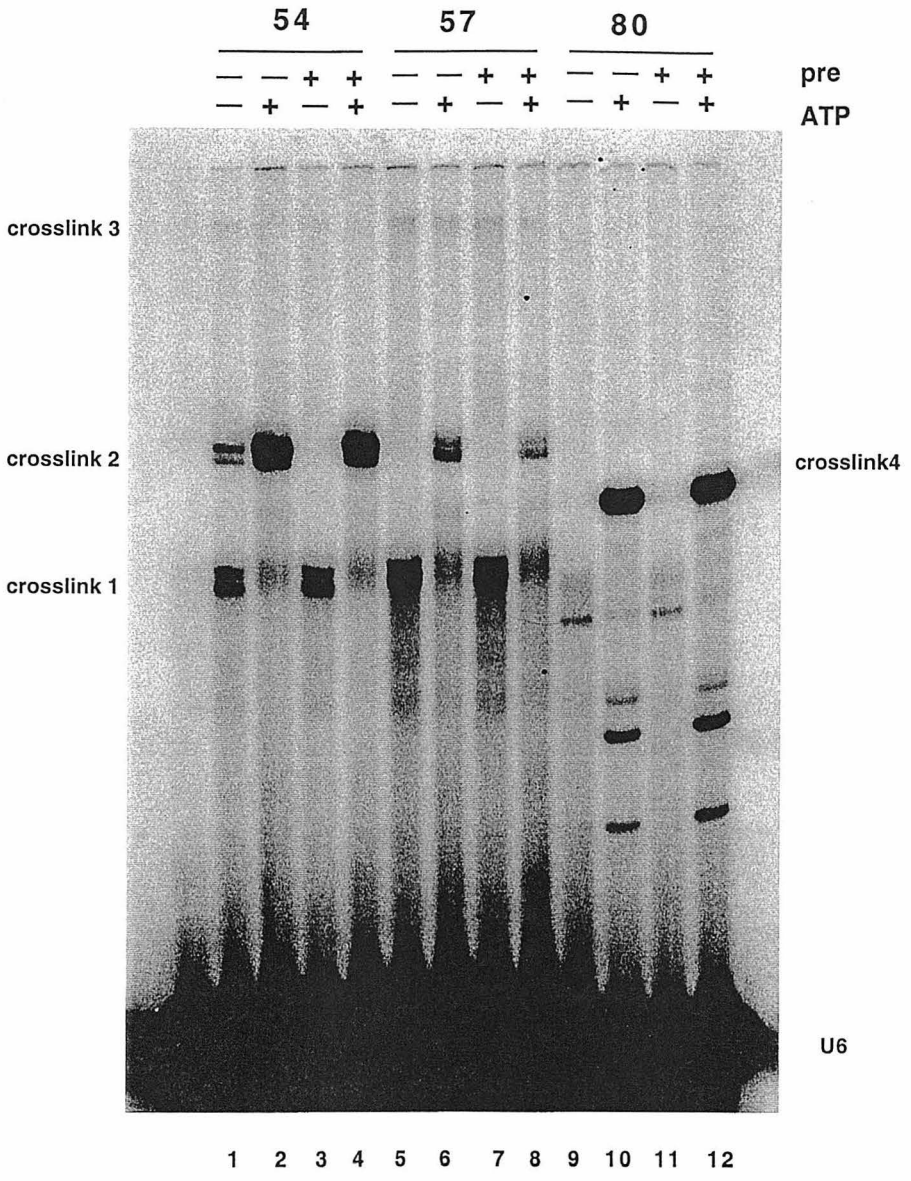
In order to form crosslinks between U6 snRNA and other RNAs in yeast spliceosomes, yeast splicing extract was depleted of endogenous U6 and reconstituted with P-^{32} labeled synthetic U6 containing a single 4-thioU. Pre-mRNA substrate was added

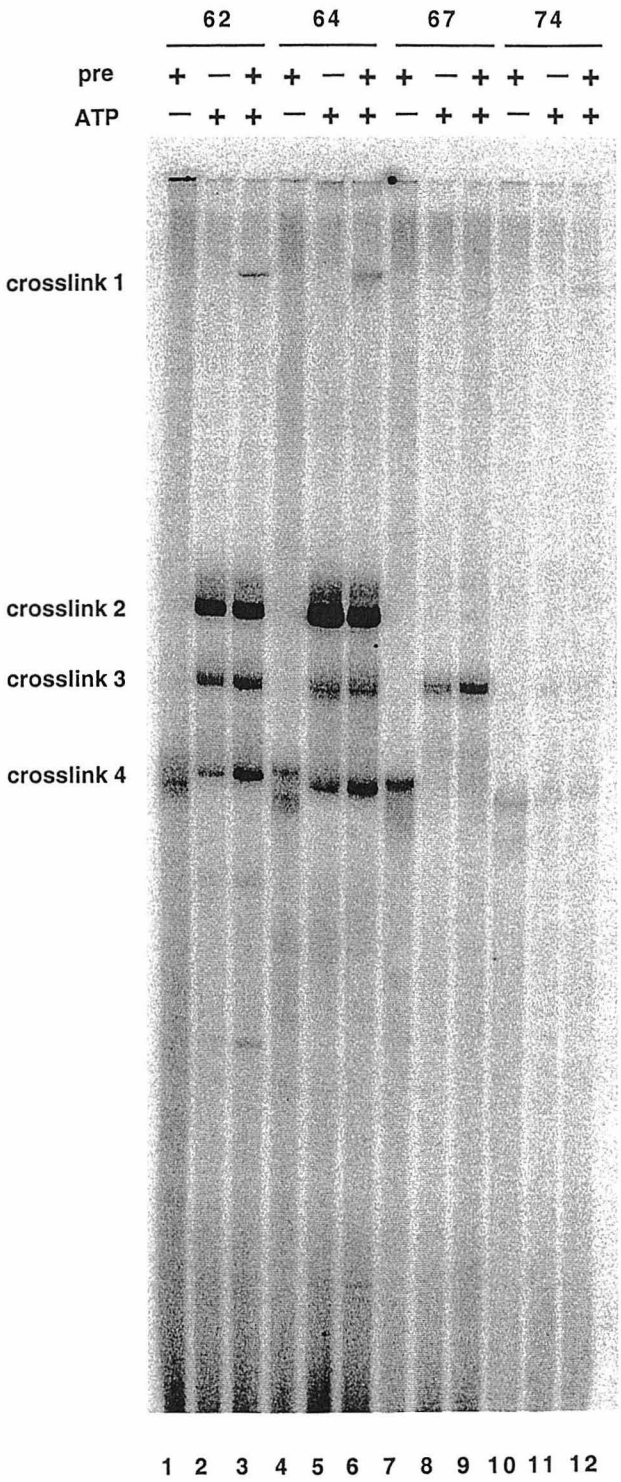
and allowed to splice for 10 min, then the sample was irradiated with long wavelength energy UV light to initiate crosslinking. In these experiments, the pre-mRNA substrate was used without the P-^{32} label, so that crosslinks to U6 RNA could be observed without interference from the degradation products of pre-mRNA. Since the pre-mRNA substrate was not labeled, it was not possible to monitor the splicing reaction as the crosslinks were formed. The presence or absence of ATP and pre-mRNA was controlled to indicate whether a crosslink is dependent on pre-mRNA splicing. A crosslink which occurs in the absence of ATP must not be dependent on splicing because ATP is required at various steps in the assembly of active spliceosomes. A crosslink which is not dependent on the addition of pre-mRNA is not dependent on splicing of the exogenous pre-mRNA. However, it is possible that there are endogenous introns in the splicing extract which use the synthetic U6 RNA, such that the low level of U6 crosslinks that we observe in the absence of added exogenous pre-mRNA might actually be dependent on endogenous pre-mRNAs. If a U6 crosslink is dependent on both ATP and added pre-mRNA, then the U6 crosslink is thought to occur in spliceosomes.

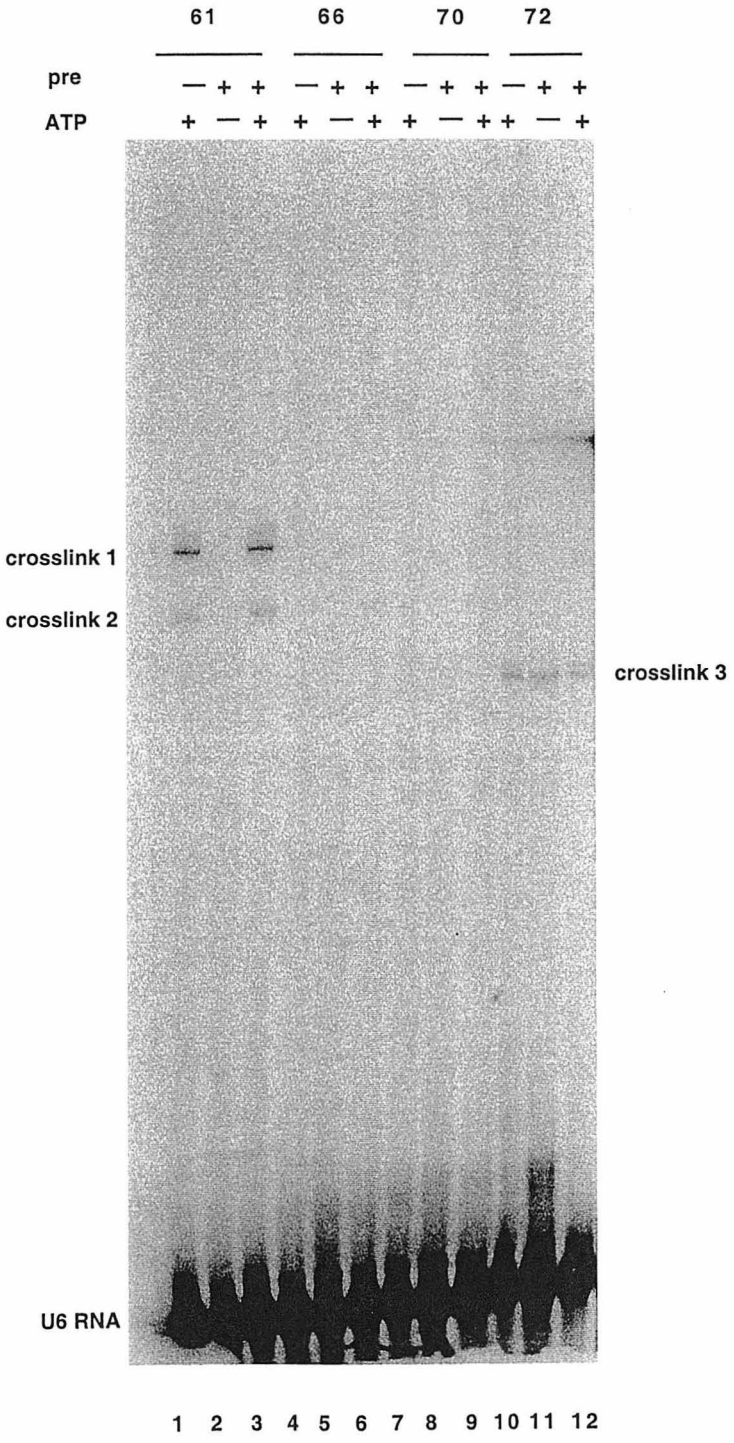
4-ThioU 51 in U6 RNA crosslinks to several RNAs in an ATP-dependent manner, and to the exogenous pre-mRNA (Fig. 2a). There is one predominant crosslinked species which forms in the presence of ATP (lanes 2 and 4). However, this crosslink is not dependent on the presence of the exogenous pre-mRNA substrate (lane 4). From the molecular weight alone, it can be inferred that the low molecular weight, high-yield, ATP-dependent crosslinks are possibly to U4 or U5, but not to a high molecular weight snRNA such as U2 or U1. It is also possible that this crosslink might be a dimer of two U6 RNAs or an intramolecular crosslink within a single U6 RNA molecule. The identity of this crosslink was subsequently established by RNAseH digestion with oligonucletides complementary to sequences in the U snRNAs (see below). In addition to this crosslink, we observed three high molecular weight crosslinks, which are dependent on the exogenous pre-mRNA (lane 2). In fact, these crosslinks are to the pre-mRNA and the

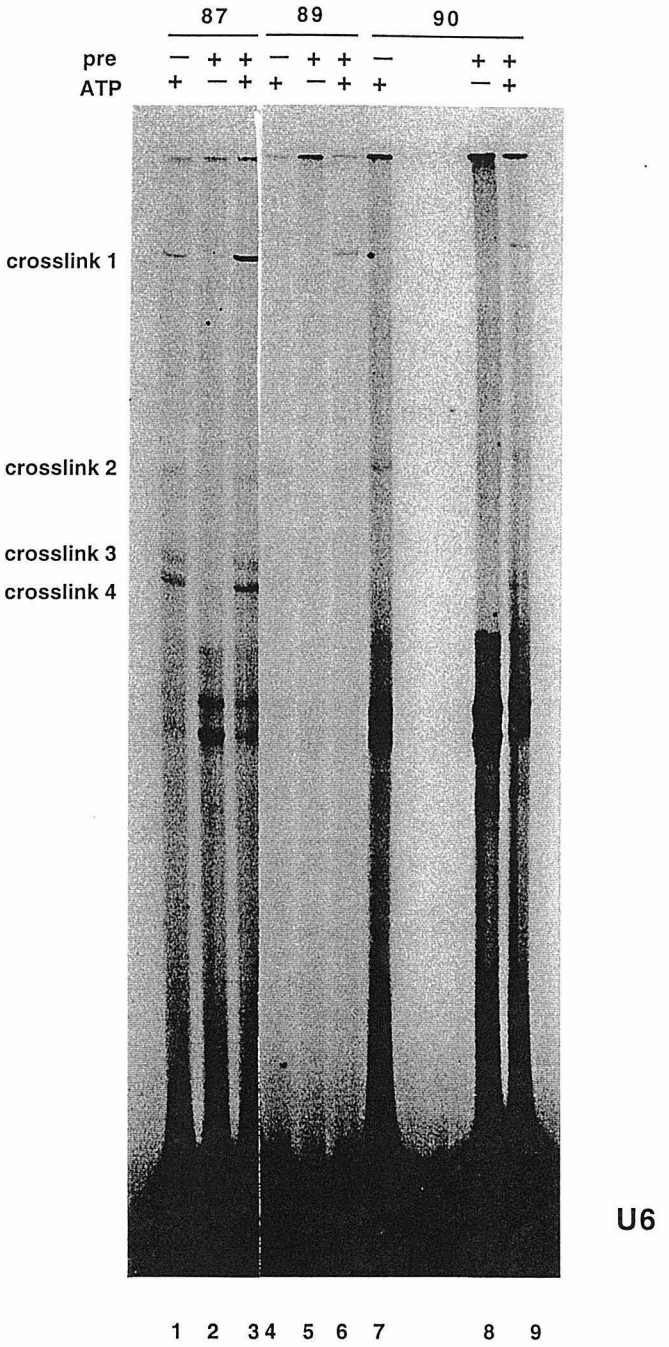
Figure. 2a, b, c,d,e. All crosslinks found to RNAs in the splicing extract using 4-thioU site-specifically substituted into U6 RNA. The U6 RNA was internally labeled with P32, but the pre-mRNA was not labeled in these experiments.


no
thioU
thioU 51
pre + + + —
ATP + + — +
(sU51)U6/
pre-mRNA
crosslinks
(sU51)U6/U4
crosslink
U6RNA
1 2 3 4


54
57
80
pre
ATP
crosslink 3
crosslink 2
crosslink 1
crosslink4
U6
1 2 3 4 5 6 7 8 9 10 11 12


62
64
67
74
pre
ATP
crosslink 1
crosslink 2
crosslink 3
crosslink 4
1 2 3 4 5 6 7 8 9 10 11 12


61
66
70
72
pre
ATP
crosslink 1
crosslink 2
crosslink 3
U6 RNA
1 2 3 4 5 6 7 8 9 10 11 12


87
89
90
pre
ATP
crosslink 1
crosslink 2
crosslink 3
crosslink 4
U6
1 2 3 4 5 6 7 8 9

lariat intermediate. In previous experiments where we crosslinked unlabeled (4-thioU51) U6 RNA to P-^{32} labeled pre-mRNA, we identified three crosslinks to the 5' splice site of the pre-mRNA or lariat intermediate (Kim and Abelson, 1996). We must be seeing these crosslinks except now with the U6 RNA P-^{32} labeled and the pre-mRNA unlabeled. There are no strong crosslinks in the absence of 4-thioU (lane 1).

Crosslinks obtained with 4-ThioU 54, 57 and 80 are shown in Fig. 2b. In the case of 4-thioU54, two bands of low molecular weight ATP-independent crosslinks are found (crosslink 1, lanes 1 and 3). Upon addition of ATP, these crosslinks diminish in intensity, and in their place, a higher molecular weight crosslink (crosslink 2) forms with higher yield. Similar results were obtained with 4-thioU at position 57, except that in this case, the higher molecular weight ATP-dependent crosslinks (crosslink 2, lane 6 and 8) appeared in lower yields than the lower molecular weight ATP-independent crosslinks (crosslink 1, lane 5 and 7). It is interesting to note that both the ATP-independent and dependent crosslinks found using 4-thioU54 co-migrate with those found using 4-thioU57. It is possible that they may be the same crosslink species, but further characterization of the crosslink was not pursued because it was not dependent on pre-mRNA. We also observed a very high molecular weight crosslink for 4-thioU57 (crosslink 3, lanes 5, 6, 7, 8), but this crosslink was independent of both ATP and pre-mRNA, and was not pursued further. Both the ATP-dependent (crosslink 2) and independent crosslinks (crosslink 1) found with 4-thioU 54 and 57 migrated too far down the gel to be crosslinks between U6 RNA and the pre-mRNA. This agrees with the negative results we obtained earlier using labeled pre-mRNA where we found no crosslinks from 4-thioU54 or 57 of U6 to the pre-mRNA substrate (Kim and Abelson, 1996) . 4-ThioU80 forms very little ATP-independent crosslinks (lanes 9 and 11), but forms low molecular weight ATP-dependent crosslinks with high yields (crosslink 4, lanes 10 and 12). The ATP-dependent crosslinks were characterized further in experiments below. Although crosslinks to the pre-mRNA were not observed in this set of crosslinking experiments using high specific activity P-^{32}

labeled 4-thioU80 U6 RNA , crosslinks to the pre-mRNA were observed in experiments with the same 4-thioU in U6 using pre-mRNA labeled to high specific activity (see below). This may be because the actin pre-mRNA was labeled to higher specific activity, thereby increasing the detection limit. Also, the (4-thioU80) U6/pre-mRNA crosslink was produced in lower yield than the (4-thioU51) U6/pre-mRNA crosslinks which were readily apparent when the U6 RNA was labeled to high specific activity and the pre-mRNA was unlabeled.

Crosslinks obtained with 4-thioU at positions 62, 64, 67, and 74 in U6 RNA are shown in Fig. 2c. The pattern of crosslinks is very similar for both 4-thioU62 (lanes 1, 2, 3) and 4-thioU64 (lanes 4, 5, and 6). Given the proximity of these two 4-thiouridines and the similarity of the crosslinking patterns obtained, it is likely that the same RNAs were crosslinked; however the crosslinked sites might differ. One striking feature is that a high molecular weight crosslink formed with 4-thioU at position 62 and 64 that was dependent on both ATP and pre-mRNA (crosslink 1, lanes 3 and 6, respectively). Positions 62 and 64 are unusual in that a 4-thioU substituent at either position forms an ATP and pre-mRNA-dependent crosslink to a spliceosomal RNA rather than to the pre-mRNA. This ATP and pre-mRNA-dependent crosslink is characterized further below. There are also 2 sets of ATP-dependent low molecular weight crosslinks (crosslink 2 and 3, lanes 2, 3, 5, 6) and a set of ATP and pre-mRNA independent crosslinks (crosslink 4, lanes 1-6) with 4-thioU 62 and 64. The crosslinks obtained with 4-thioU67 and 74 are much lower in yield. The two 4-thioU's appear to have completely similar crosslinking patterns to each other and partially similar crosslinking patterns to those obtained with 4-thioU 62 and 64. In both 67 and 74, there aare lower yields of low molecular weight crosslinks (crosslinks 3 and 4, lanes 8, 9 and 11, 12) which co-migrate with those found with 4-thioU62 and 64. There is also a faint trace of the ATP and pre-mRNA dependent high molecular weight crosslink (crosslink 1: lane 9 for 4-thioU67, and lane 12 for 4-thioU74), but it appears to migrate much faster on the gel than the corresponding crosslink found for 4-thioU62 and 64.

However, crosslink 2 which occurs with high yield for 4-thioU 62 and 64 are markedly absent for 4-thioU 67 and 74.

In Fig. 2d, the crosslinks found with 4-thioU 61, 66, 70 and 72 are shown. Few crosslinks were obtained with this series of 4-thiouridines. 4-ThioU61 gave a pair of ATP-dependent crosslinks -- one dark band (crosslink 1), and one lighter band below it (crosslink 2) -- (lanes 1 and 3), while 4-thioU 72 gave a single crosslink which appeared in all lanes (crosslink 3, lanes 10,11,12). This latter crosslink was clearly not dependent on splicing because it was independent of both ATP and pre-mRNA. 4-ThioU 66 and 70 curiously did not yield any RNA crosslinks at all.

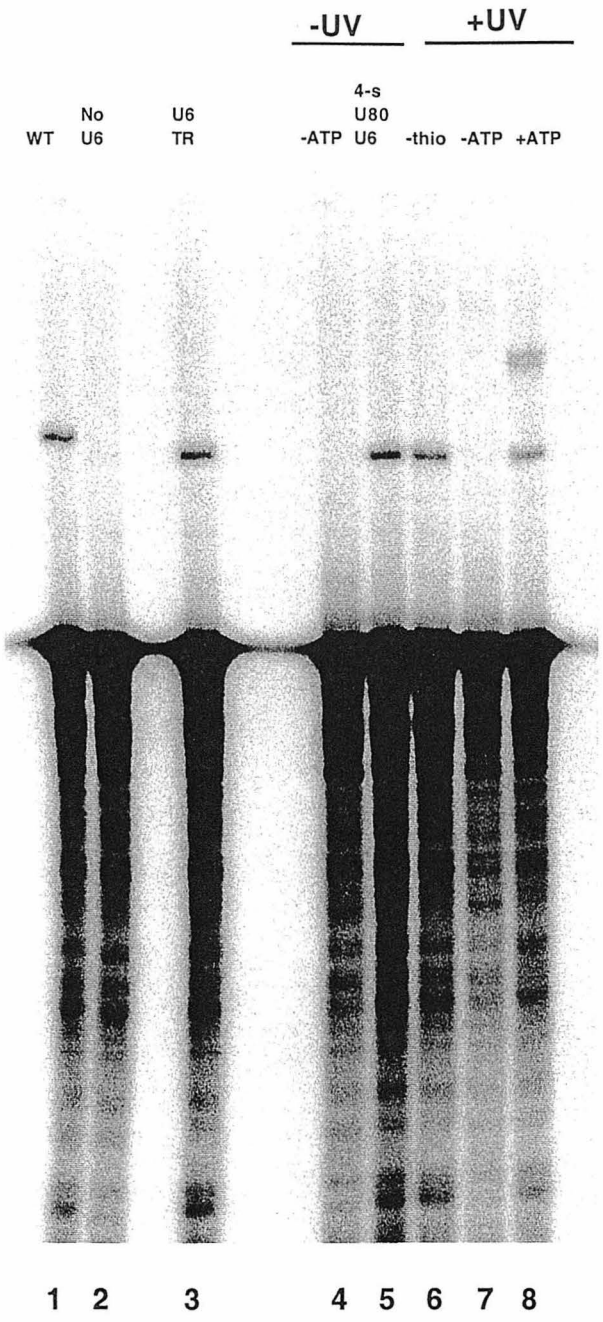
The last panel for high specific activity 4-thioU U6 RNA crosslinks shows results for 4-thioU at positions 87, 89 and 90 (refer to Fig. 2e). We observed an ATP-dependent and pre-mRNA enhanced crosslink in high yield for 4-thioU87 (the top band (crosslink 1) in lane 3). The enhancement in the presence of exogenous pre-mRNA is approximately 20 fold as quantitated by a Molecular Dynamics Phosphorimager. The low amount of this crosslink produced in the absence of added pre-mRNA may depend on the presence of endogenous pre-mRNAs in the yeast extract. If so, the crosslink would be strictly dependent on pre-mRNA splicing. The same crosslink appears to be produced by 4-thioU89 and 90, although in these two cases, the yields of the crosslink are greatly diminished. In fact, for these two 4-thioU's, no trace of the crosslink is observed in the minus pre-mRNA lane (lane 6 and 9). This may be because the yield of the crosslink is diminished to the point where it cannot be visualized. Three lower molecular weight ATP-dependent crosslinks were formed by irradiation of (4-thioU87) U6 RNA (crosslink 2, 3, 4 : lanes 1 and 3). Only crosslink 2 forms with 4-thioU89 and 90, even though this crosslink has the lowest yield of the three (cf. lanes 1 and 3 with lanes 4 and 6 or with lanes 7 and 9). Below these ATP-dependent bands, there are three ATP and pre-mRNA independent bands.

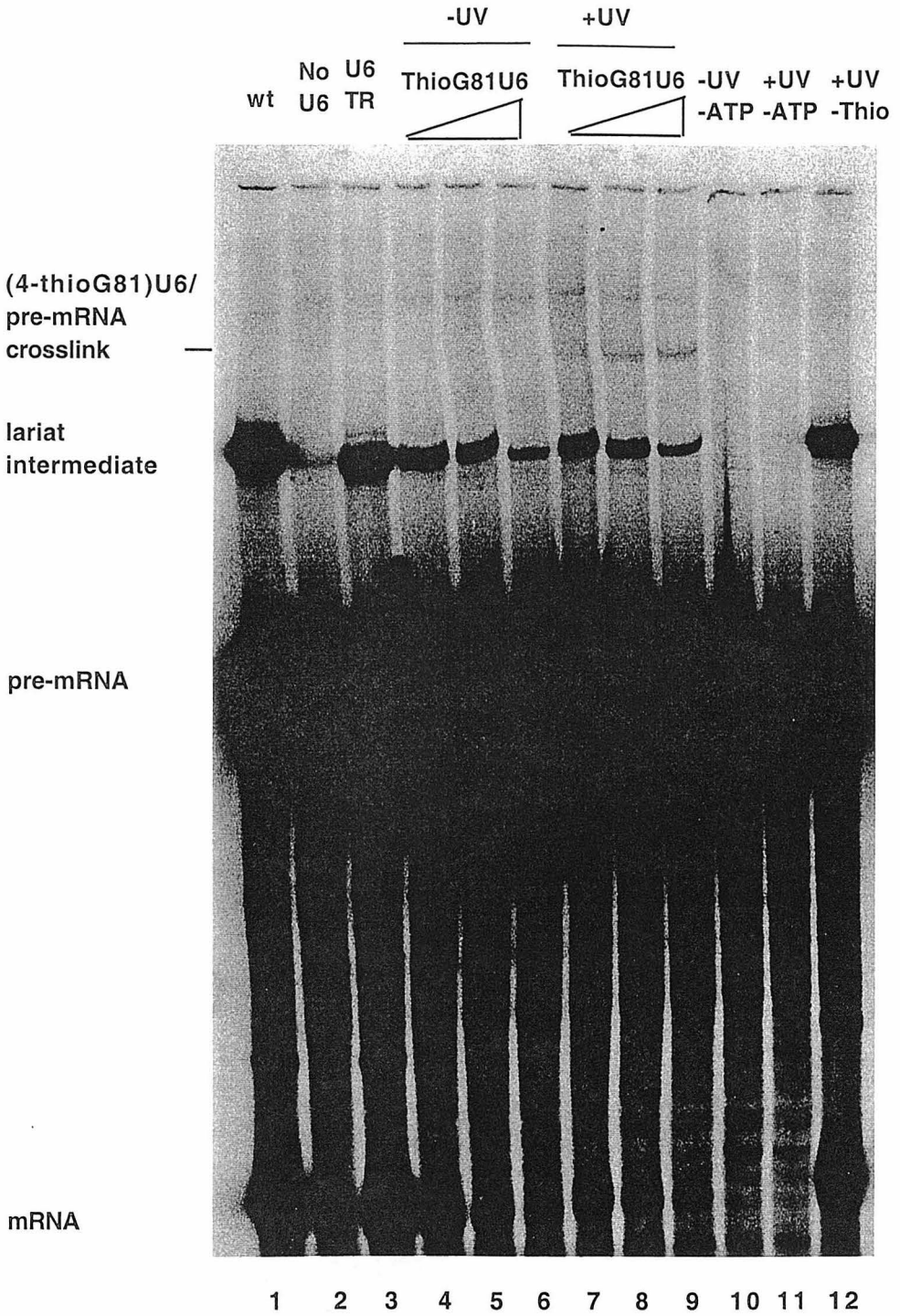
Crosslinks of 4-thioU80, 6-thioG 81, 55 U6 RNA with actin pre-mRNA

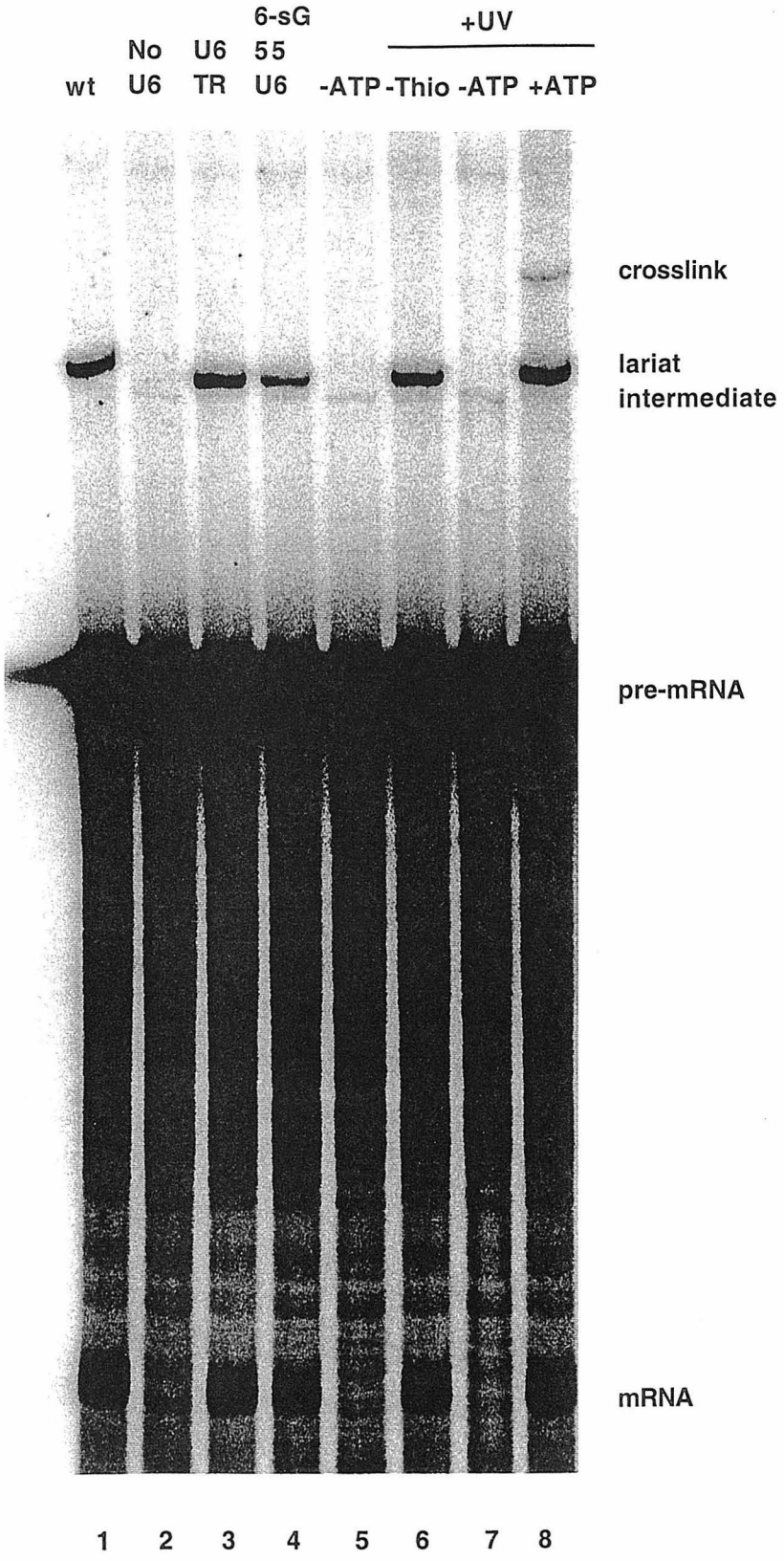
Crosslinks between 4-thioU-substituted U6 RNA and pre-mRNA labeled to high specific activity are shown in Fig. 3 a, b, and c. In these experiments the pre-mRNA substrate is labeled to much higher specific activity than the U6 RNAs such that the radioactivity of the U6 RNA band is "hidden" by the strong radioactive bands resulting from the numerous degradation products of the actin pre-mRNA (notice the ladder of bands below the pre-mRNA in each lane of Fig. 3a). However, since, the pre-mRNA is P-^{32} labeled, we can monitor the progress of the splicing reaction concurrently with the formation of a crosslink.

The crosslinks produced between 4-thioU80 in U6 RNA and actin pre-mRNA are shown in the first panel, Fig. 3a. In the wild-type control lane (lane 1), the in vitro splicing activity of a yeast splicing extract is shown. The internally P-^{32} labeled actin pre-mRNA was efficiently spliced and the lariat intermediate was apparent as a distinct band above the pre-mRNA. On this 5% polyacrylamide denaturing gel, a product of the second step of splicing, the lariat intron, co-migrates with the pre-mRNA band, and thus cannot be visualized. However, on higer percentage gels, we observed that the lariat intron formed (data not shown). The other product of the second step of splicing, namely the spliced mRNA is also not visible in Fig. 3a because of the high background from the degradation products of the actin pre-mRNA substrate. However, in Fig. 3b and c the spliced messenger RNA can be more distinctively visualized. When the splicing extract is incubated with the deoxynucleotide (d1) complementary in sequence to a central portion of U6 RNA, in vitro splicing activity is abolished completely because the RNAseH activity degrades the endogenous U6 RNA (Fabrizio et al., 1989). The d1 oligonucleotide itself is degraded by the DNAses in the extract upon further incubation, so that synthetic U6 RNA added later is not cleaved and the splicing activity restored (lane 3, Fig. 3a). Synthetic U6 RNA ligated from three pieces (lane 6), and containing 4-thioU at position 80 (lane 7) can

Figure. 3a, b, c Crosslinks to pre-mRNA substrate found using 4-thioU or 6-thioG site-specifically incorporated into U6 RNA. In these experiments, the actin pre-mRNA substrate was internally labeled with P^{32}, so that the reconstitution of splicing activity and the course of the splicing reaction could be followed.


-UV
+UV
WT
No U6
U6 TR
-ATP
4-s U80 U6
-thio
-ATP
+ATP
1 2 3 4 5 6 7 8


-UV
+UV
wt
No U6
U6 TR
ThioG81U6
ThioG81U6
-UV -ATP
+UV -ATP
+UV -Thio
(4-thioG81)U6/ pre-mRNA crosslink
lariat intermediate
pre-mRNA
mRNA
1 2 3 4 5 6 7 8 9 10 11 12


wt
No U6
U6 TR
6-sG 55 U6
-ATP
+UV
-Thio
-ATP
+ATP
crosslink
lariat intermediate
pre-mRNA
mRNA
1 2 3 4 5 6 7 8

also restore the splicing activity of the U6-depleted extract. Irradiation of the U6-reconstituted extract with long wave UV generated a high molecular crosslink in the presence of ATP, 4-thioU and UV irradiation (lane 8). This crosslink does not appear in the absence of ATP (lane 7), 4-thioU (lane 6) or UV irradiation (lane 5). Lane 4 shows that splicing activity is abolished when extracts are depleted of ATP via glucose addition. The crosslink shown in lane 8 is produced in good yield -- 50% of the partially spliced lariat intermediate. The crosslink was excised from the gel, electro-eluted and used as a template for primer extension to map the site of the crosslink on the pre-mRNA (see the Materials and Methods section above).

In Fig. 3b, we present the results of crosslinks found with 6-thioG at position 81 in U6 RNA with the pre-mRNA. As in Fig. 3a, we can observe splicing of the P-^{32} labeled actin pre-mRNA in the wild-type extract (lane 1). Splicing activity was abolished in the U6-depleted extract (lane 2) and restored upon addition of synthetic U6 RNA transcript (lane 3) or increasing amounts of 6-thioG 81 U6 RNA (lanes 4, 5, and 6). When the U6-complemented extract was irradiated with UV light, a crosslink was generated in the presence of ATP (lanes 7, 8, and 9), but not in the absence of ATP (cf.- ATP lane 11) or 6-thioG (cf. -ThioG lane 12). Lane 10 shows that depletion of ATP in the extract results in the loss of splicing activity.

Fig. 3c shows crosslinks found between 6-thioG at position 55 in U6 RNA and the pre-mRNA. Reconstitution of splicing activity from U6 depleted extract (lane 2) by U6 RNA transcript (lane 3) and by ligated U6 RNA with 6-thioG at position 55 (lane 4) are shown in the first four lanes. Depletion of ATP results in loss of splicing activity (lane 5). In the presence of ATP (cf. -ATP lane 7) and 6-thioG81 (cf. -thio lane 6), U6 RNA forms a crosslink with the P-^{32} labeled splicing substrate upon UV irradiation (lane 8).

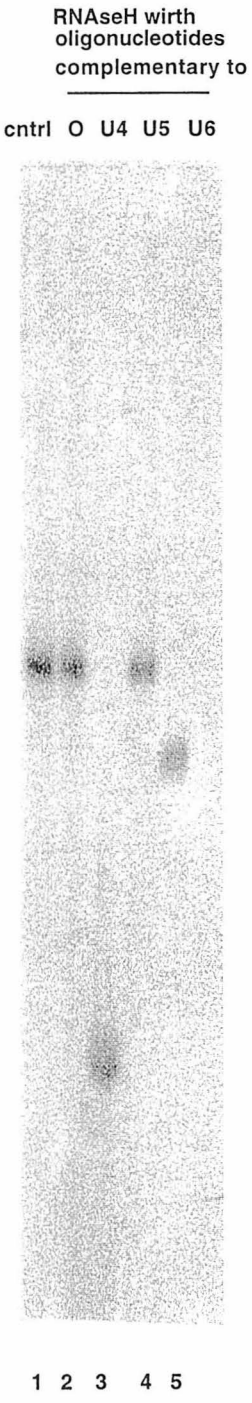
4-ThioU51 and 4-thioU 80 in U6 RNA crosslink to U4 snRN

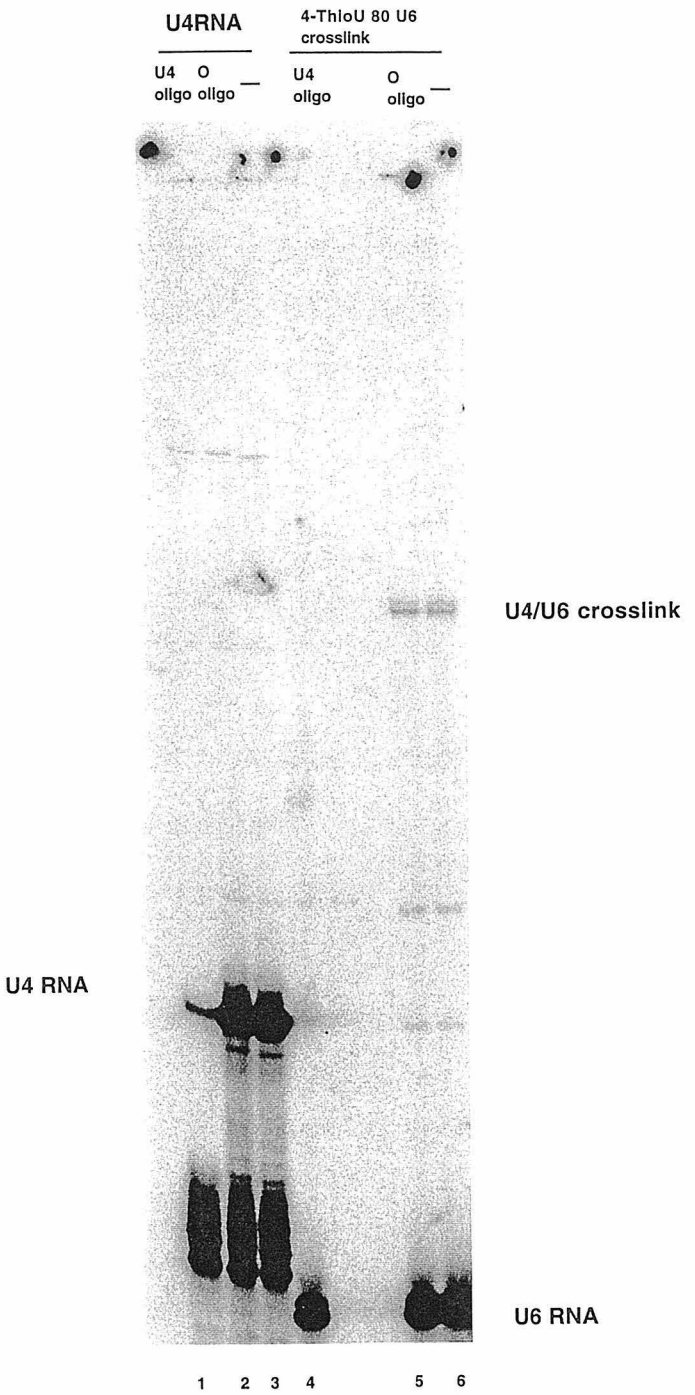
4-thioU 62 and 64 crosslinks to U2 snRNA

As observed in the panel of data shown in Fig. 2, many ATP-dependent crosslinks formed between 4-thioU in U6 RNA and a low molecular weight spliceosomal RNA. In two of these cases (for 4-thioU51 and 4-thioU80), the crosslinked spliceosomal RNA were identified by oligonucleotide-directed RNAseH degradation using oligonucleotides complementary to the low molecular weight snRNAs -- namely U4 or U5 snRNA (see Fig. 4 a and b). In the control lane (lane 1) of Fig. 4a, we can see the (4-thioU51) U6 RNA crosslink to the low molecular weight spliceosomal RNA. This crosslink is not degraded during incubation with RNAse H in the absence of any oligonucleotide (lane 2). Therefore, the RNAseH enzyme does not cleave the crosslink in the absence of added oligonucleotide. However, the crosslink is efficiently degraded when incubated with RNAse H in the presence of DNA oligonucleotides complementary to U4 (lane 3) and U6 (lane 5), but not to U5 (lane 4). This demonstrates that the crosslink is between U4 snRNA and 4-thioU51 of U6 RNA.

Using the same RNAseH assays, we also show that 4-thioU80 of U6 RNA forms an ATP-dependent low molecular weight crosslink to U4 snRNA (Fig. 4b). In these experiments, a control digestion of a U4 RNA transcript is shown in the first three lanes. The U4 RNA transcript is sensitive to RNAseH degradation when incubated with an oligonucleotide complementary to U4 (lane 1). Lane 3 shows the untreated U4 RNA transcript. In the absence of any oligonucleotide, no RNAseH degradation occurs (lane 2). In the next three lanes, the crosslink generated for 4-thioU80 U6 RNA was treated with RNAseH. In the presence of a U4-complementary deoxyoligonucleotide, the crosslink is efficiently degraded (lane 4). Lane 6 shows the untreated U4/ U6 RNA crosslink. In the absence of any oligonucleotide, the crosslink is not degraded during incubation with RNAseH.

Figure. 4 a, b 4-thioU 51 or 80 in U6 snRNA forms highly efficient ATP-independent crosslinks with U4 snRNA. One of the RNA species crosslinked to the 4-thioU in the P32-labeled U6 RNA was identified by RNAseH digestion with oligonucleotides complementary to U4 snRNA. Controls include RNAseH digestion of the crosslink in the absence of any oligonucleotides and untreated crosslink.


RNAseH wirth
oligonucleotides
complementary to
cntrl O U4 U5 U6
1 2 3 4 5


U4RNA
4-ThioU 80 U6 crosslink
U4 oligo
O oligo
U4 oligo
O oligo
U4/U6 crosslink
U4 RNA
U6 RNA
1 2 3 4 5 6

The RNA which forms an ATP and pre-mRNA dependent crosslink to 4-thioU64 was identified by RNAseH digestion with oligonucleotides complementary to sequences in the snRNA and the pre-mRNA (Fig. 5). Lane 1 shows that the faint high molecular weight crosslink is not degraded when incubated with RNAseH in the absence of any oligonucleotide. The crosslink is sensitive to RNAseH digestion in the presence of oligonucleotides complementary to two different sequences in U2 (lane 3 and 4). This crosslink is insensitive to RNAseH digestion with oligonucleotides complementary to sequences in U1 (lane 2), U4 (lane 5), U5 (lane 6), the 5' exon (lane 7) or the intron (lane 9). These results show that the crosslink of 4-thioU64 U6 RNA is to U2 snRNA, and not to any other snRNA or the pre-mRNA.

4-thioU80 in U6 RNA crosslinks to A210 in the intron, upstream of the branchpoint

The site of the (4-thioU80) U6 crosslink on the pre-mRNA, discussed above (Fig. 3a), was mapped by primer extension analysis (Fig. 6). A sufficient quantity (0.1 fmole) of the crosslink was recovered by electro-elution from gel-slice and used as the RNA template for primer extension. A P-^{32} labeled primer complementary in sequence to the 3' exon (called 3'F) was extended using AMV reverse transcriptase. The primer extension is blocked one nucleotide 3' of the crosslinked site because the crosslink acts as a physical barrier for the reverse transcriptase. The precise nucleotide on the pre-mRNA which was crosslinked to U6 RNA formed a crosslink was determined by comparing the primer extension stop with dideoxy-sequencing bands (lanes 1-4, Fig. 6) for the pre-mRNA gene using the same primer (i.e. 3'F). The primer extension on the gel-purified crosslink was strongly blocked at a single site (lane 6). This single band was matched with pre-mRNA nucleotide A211 in the sequencing lane, which means that the site of the crosslink was one nucleotide upstream of this nucleotide, namely the A210 in the intron. The fact that the

Figure.5 4-thioU introduced at position 62 or 64 (only the data for 64 crosslink is shown here) in U6 snRNA forms crosslinks with U2 snRNA in an absolutely pre-mRNA and ATP-dependent manner, suggesting that they are splicing-dependent (refer to Fig. 2a). The RNA crosslinked to the 4-thioU 62 or 64 in U6 was identified by RNAseH digestion of the crosslinks with oligonucleotides complementary to the U snRNAs (U1, U2, U4, and U5) or to sequences in the pre-mRNA (5' exon and the intron). The high molecular weight crosslink is only sensitive to RNAseH digestion using two oligonucleotides complementary to sequences in the U2 snRNA.

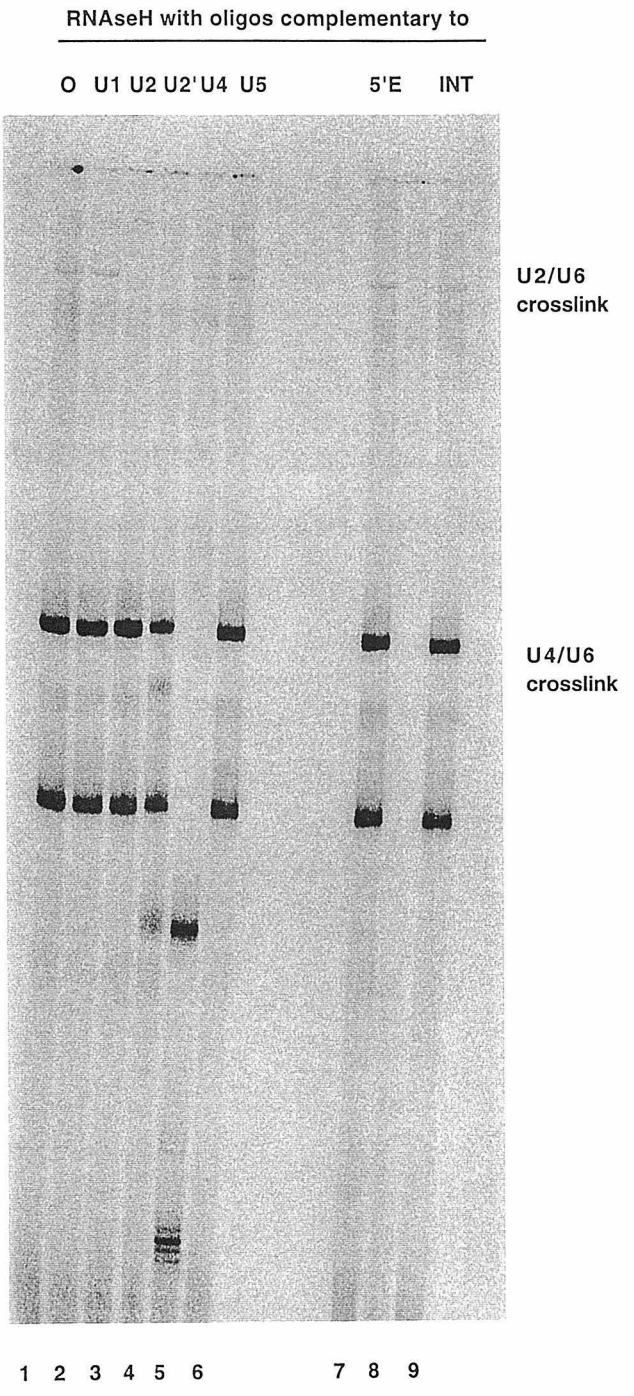
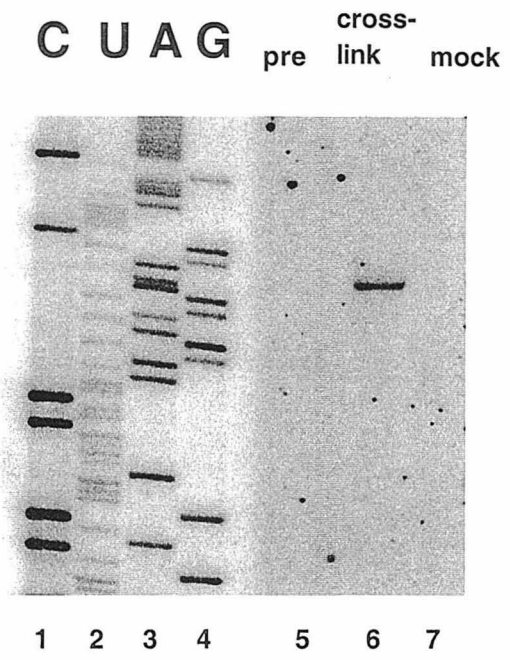

RNAseH with oligos complementary to
O U1 U2 U2' U4 U5 5'E INT
U2/U6 crosslink
U4/U6 crosslink
1 2 3 4 5 6 7 8 9

Figure. 6 The site of the U80 U6 snRNA crosslink to the pre-mRNA was mapped by primer extension to A 210 in the yeast actin intron sequence. There is a single strong stop at the site of the crosslink. No primer extension stops are observed in the control lanes using pre-mRNA as or mock UV-irradiated extract RNA as template for primer extension.


C U A G
pre
cross-
link
mock
1 2 3 4 5 6 7

primer extension using oligonucleotide 3'F was not blocked prematurely at the branchpoint suggests that the crosslink was to the pre-mRNA species and not to a lariat intron intermediate. We confirmed that extension of the primer on a lariat intron species is blocked at the branchpoint (data not shown). In a control, primer extension using UV-irradiated pre-mRNA that was gel-purified from a duplicate sample showed no primer extension stops (lane 5) that coincided with those obtained using the gel-purified crosslink. This suggests that there were no strong secondary structures or UV nicks in the pre-mRNA which caused the stop found for the crosslinked pre-mRNA. A mock control with total extract RNA co-migrating with the 4-thioU80 crosslink (lane 7) was done to rule out the possibility that the A210 stop could be caused by RNAs in the extract which "fortuitously" co-migrated with the crosslink.

Discussion

We have done an extensive search for splicing-dependent crosslinks to 4-thioU site-specifically incorporated into several positions in U6 RNA (refer to Table 1 for a summary). We are able to look for crosslinks between U6 snRNA and other spliceosomal snRNAs because Fabrizio et al. in our laboratory had developed an in vitro system for reconstituting U6 snRNPs using synthetic U6 RNA(Fabrizio et al., 1989). In the mammalian system, where U6 cannot be reconstituted in vitro, it was only possible to incorporate 4-thioU in the pre-mRNA (Sontheimer and Steitz, 1993).

In a few of the U6 sites tested (66, 67, 70, 72, 74), a 4-thioU substituent does not form any crosslinks to RNAs. It is interesting that these nucleotides are all in the Brow stem intramolecular helix. As we have discussed earlier, it is possible that proteins interacting with this helix can block the formation of crosslinks to RNAs. Nucleotide C66 is a strong DMS protection site in glycerol gradient purified spliceosomes (S. L. Yean and R. J. Lin, personal communication). We have observed crosslinks to unidentified proteins at this site (data not shown). Positions 70, 72 and 74 are located in the loop of the Brow stem helix -- possibly a recognition element of an RNA-binding protein. We have also previously observed 2'-deoxy substitutions which block splicing at positions 70 and 72 (Kim et al., 1997). The 2'-OH's at these positions may thus be engaged in a crucial interaction with a protein. In cases where we apparently obtained no crosslinks, we cannot rule out that a 4-thioU substitutent at these positions may interact with another RNA at sites that are not conducive to crosslink formation.

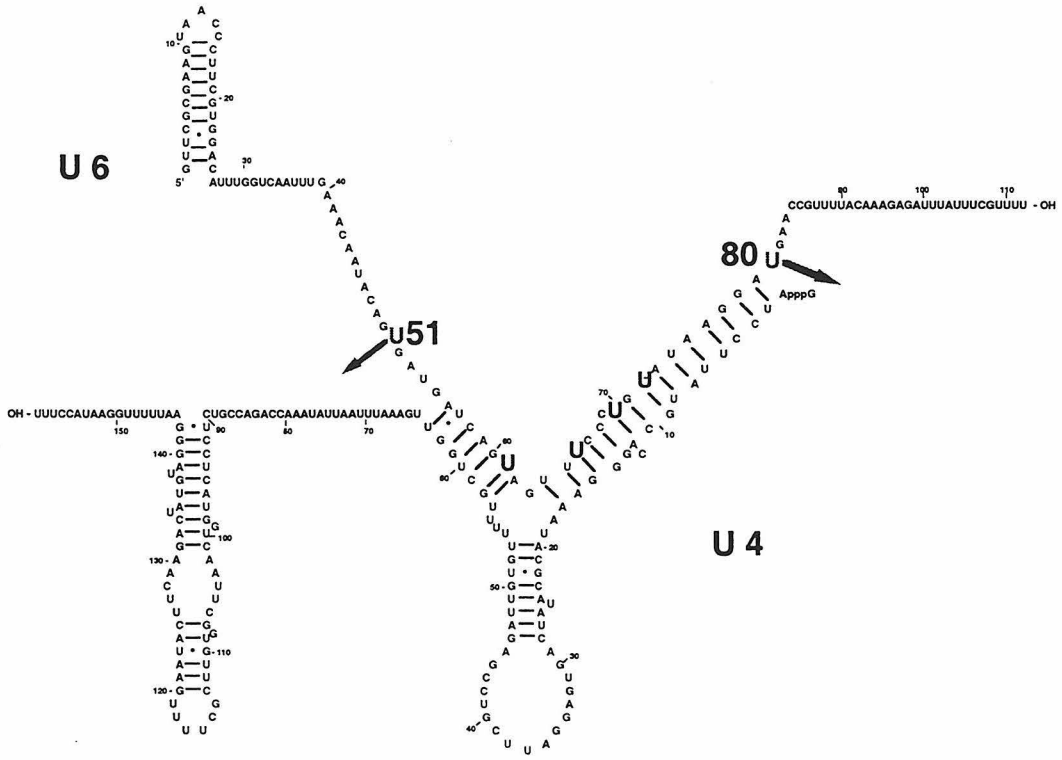
Table 1. Summary of all crosslinks obtained.

4-thioU position	crosslink partner	ATP(+/-)	pre-mRNA (+/-)	crosslinked site
46	pre-mRNA	+	+	5' splice site
51	pre-mRNA	+	+	5' splice site
	lariat intermediate	+	+	
	U4	-	-	
54	U4	+	-	
55*	pre-mRNA	+	+	
57	U4	+	-	
61	U4			
62, 64	U2	+	+	
66	proteins	-		
67	-			
70	-			
72	-			
74	-			
80	pre-mRNA	+	+	intron 210
	U4	+	-	
81	pre-mRNA			
89				
90				

Surprisingly, some crosslinks to U6 RNA were not dependent on pre-mRNA and ATP. There are many low molecular weight crosslinks apparent in all control lanes, and a high molecular weight band for 4-thioU57 which forms in the absence of ATP and pre-mRNA. These crosslinks probably form in U6 snRNPs that were not incorporated into functional spliceosomes. Clearly, the incorporation of U6 RNA into spliceosomes is much less than 100%, especially because we add an excess of the synthetic U6 RNA (50 fmoles in a 5 μl splicing reaction) in order to achieve functional U6 snRNP reconstitution.

The most efficient crosslinks obtained were between U6 and U4 (characterized for 4-thioU51 and 80) (Fig. 7a). These were ATP-dependent, but were not dependent on the presence of exogenous pre-mRNA. Presumably, U4/U6 crosslink are formed early in spliceosome assembly, prior to the dissociation of U4 from the spliceosome, possibly in the U4/U6 snRNPs or in the U4/U6.U5 tri-snRNPs which form in the absence of pre-mRNA. The U4/U6 base-pairing is known to occur in both U4/U6 snRNPs and U4/U6.U5 snRNPs, but U4 dissociates in the context of the spliceosome (Moore et al., 1993). The fact that the U4/U6 crosslinks are formed in high yield is not surprising since the association between U6 and U4 RNA is extensive, and involves 24 base-pairs in yeast (Fabrizio et al., 1989; Madhani et al., 1990). Interestingly, 4-thioU51 and 80 which crosslink to U4 RNA are outside this extensively base-paired region (4-thioU cannot form crosslinks between Watson-Crick base-pairs (Favre, 1990)). This is consistent with the fact that 4-ThioU does not form crosslinks if it is at an internal position in a base-paired sequence. Although we only characterized U4-U6 crosslinks for 4-thioU51 and 80 substituted U6, it is likely that other 4-thioU positions in our screen also crosslinked to U4 since many of these experiments showed high yield, ATP-dependent crosslinks that approximately co-migrated with the characterized U4/U6 RNA crosslinks. In this regard, it is very gratifying to discover that these ATP-dependent, putative U4/U6 crosslinks are found only with 4-thioU's which are not buried in an extensively Watson-Crick base-paired region (Fig. 7a: 4-thioU 51, 54, 57, 62, 64, 87, 89, and 90), and not found with 4-

Figure 7A. U4/U6 crosslinks found with 4-thioU at position 51 and 80 in U6 RNA. These crosslinks are likely to occur prior to the crosslinks to pre-mRNA from 4-thioU at the same positions in the fully assembled catalytically active spliceosome (Fig. 7b) since U4 is dissociated from the spliceosome prior to the first chemical step.

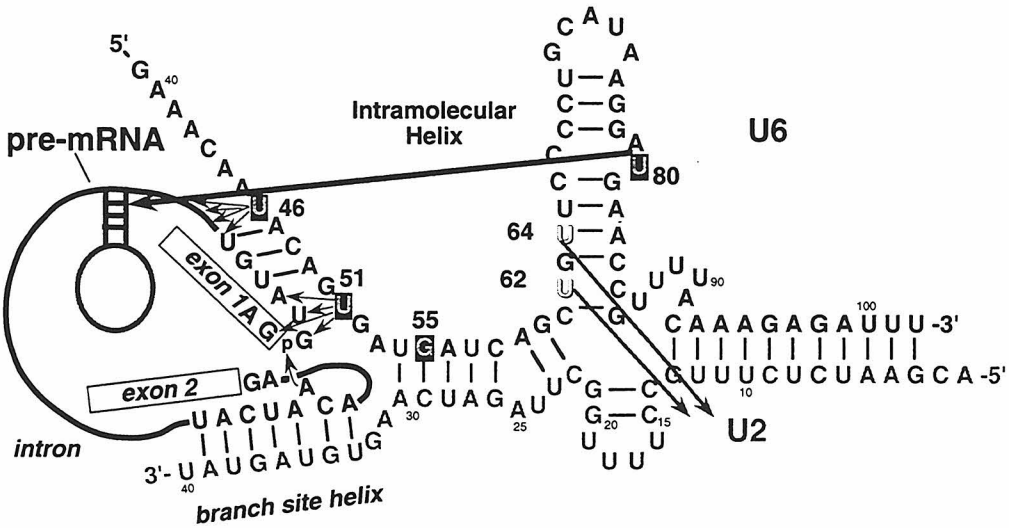

U 6
U51
80 U
ApppG
U 4
OH - UUUCCAUAAGGUUUUUAA
CUGCCAGACCAAAUAUUAAUUUAAAGU
CCGUUUUACAAAGAGAUUUAUUUCGUUUU - OH
5'
AUUUGGUCAAUUU G

thioU's that are embedded in the U4/U6 helix (Fig. 7a: 4-thioU 66, 67, and 74). Since the characterized U4/U6 RNA crosslinks occur in such high yields, they can serve as biochemical markers for the interaction between U4 and U6 RNA. For example, it would be interesting to see if the U4/U6 crosslinks occurred in ts-mutant extracts (e.g. prp24) depleted of U4/U6 re-annealing activity.

We also found crosslinks from 4-thioU 62 and 64 to U2 snRNA. These crosslinks are strictly dependent on ATP and pre-mRNA, suggesting that they occur in functional spliceosomes. However, the yield is much lower than for U4/U6 crosslinks. It is possible that the reactive conformation of the crosslinking reaction is difficult to achieve, resulting in lower yields of the U2-U6 crosslinks. At any rate, we were unable to recover sufficient crosslinked material via gel purification and eletro-elution to use as RNA template in primer extension mapping experiments. Nonetheless, it is interesting that the nucleotides 62 and 64 can crosslink to U2 RNA, since they are already base-paired to nucleotides at the 3'-end of U6 as part of the 3'-stem loop (Fig. 7b). These nucleotides can form the crosslink to U2 snRNA when the 3'-stem loop in U6 has not formed, or they can crosslink to U2 as part of the 3' stem-loop nucleotides. Since nucleotides 62 and 64 are also base-paired to U4 RNA (Fig. 7a), it is unlikely that the two nucleotides are ever free to crosslink to U2 RNA directly. Moreover, since the 4-thioU 62 and 64 crosslinks are very strongly ATP and pre-mRNA dependent, it is likely that the crosslinks to U2 occur either late in spliceosome assembly or during the catalytic steps of splicing.

In crosslinking experiments using actin pre-mRNA radioactively labeled to high specific activity, we found ATP and pre-mRNA-dependent crosslinks between pre-mRNA and U6 substituted with 6-thioG at U6 positions 55 and 81. Genetic experiments have shown that G55 of U6 is within U2/U6 helix 1a (see Fig. 1), so it is surprising that the same G55 also interacts with the pre-mRNA. When 6-thioG of U6 is base-paired to U2, the 6-thiocarbonyl group is not engaged in hydrogen bonding, and it could conceivably be

Figure 7B. The site where U80 forms a crosslink is indispensable for splicing in vitro and may be part of an intronic enhancer element which has the potential for base-pairing two distant regions of the intron, thereby "pinching" the intron close together and shortening the intron. This is the first time that U80 in U6 snRNA has been shown to interact with any other component in the spliceosome. Phosphorothioate substitution 5' to this U80 blocks splicing in yeast and nematodes, and also in domain V of Group II introns. The 4-thioU 46 and 51 crosslink to the 5' splice site were reported in previous work (29).


5'
pre-mRNA
Intramolecular
Helix
U6
U2
exon 1
exon 2
intron
branch site helix
3'-
-3'
-5'

available to the pre-mRNA for crosslinking. Alternatively, the U6 interaction with the pre-mRNA could occur at a different time than the interaction with U2 snRNA. Unfortunately, all of the 6-thioG crosslinks were unstable and could not be recovered intact for mapping by primer extension.

We were more successful in characterizing the crosslinks to the pre-mRNA obtained for (4-thioU80) U6 RNA (7b). The conserved bulged U80 in the U6 intramolecular 3' stem-loop crosslinks to the pre-mRNA in an ATP and pre-mRNA-dependent manner, suggesting that it is splicing dependent. We were able to map the site of this crosslink on the pre-mRNA by primer extension analysis. The 4-thioU80 position in U6 RNA was crosslinked to the actin pre-mRNA intron at position 210, which is 56 nucleotides upstream of the branchpoint. Nucleotide A210 is located in the actin intron upstream of the branchpoint in a region which is essential for in vitro splicing(Newman, 1987). The nucleotides in this region have the potential to base-pair with intronic sequences just downstream of the 5' splice site. This base-pairing would effectively bring the 5' splice site much closer to the branch site. Such elements are found in the pre-mRNAs of at least three other yeast genes -- CYH2, rp28 and rp51 (Charpentier and Rosbash, 1996; Goguel and Rosbash, 1993; Libri et al., 1995; Newman et al., 1985). Similar pairing interactions, in similar locations, can be drawn for many large yeast introns (Parker and Patterson, 1987). In the CYH2 gene, the intramolecular base-pairing partners are found between a sequence just downstream of the 5' splice site and a complementary sequence upstream of the branchpoint (Newman, 1987). Deletion of the element upstream of the branchpoint abolishes splicing in vitro and in vivo. Deletion of the element downstream of the 5' splice site does not block splicing, but rescues splicing of pre-mRNA lacking the element upstream of the branchpoint. From this data, a model was suggested whereby the 5' complementary element forms an inhibitory purine/pyrimidine base pair within the element itself, which is relieved by base-pairing with the 3' complementary element or by deleting the 5' element itself. In the rp51b gene, both the 5' complementary element and the 3'

complementary element are needed for efficient splicing in vivo (Goguel and Rosbash, 1993; Libri et al., 1995) and in vitro (Charpentier and Rosbash, 1996). In fact, randomization and selection experiments have shown that this base-pairing element is a general feature that is required for rp51b intron splicing (Libri et al., 1995), and no other required intronic structure was found in this pre-mRNA. This base-pairing between the intron downstream of the 5' splice site with intron sequences upstream of the branchpoint could facilitate the first step of splicing, and was termed an "intron enhancer". The intron enhancer may be complementing the interaction between U6 and 5' splice site, and that between U2 and the branchpoint to bring the branchpoint close to the 5' splice site for the first step of splicing. Thus, it is not difficult to imagine that U6 RNA, which is known to interact with the 5' splice site, becomes crosslinked to intron enhancer sequences near the branchpoint sequence. Paradoxically, in vitro data showed that the intron enhancer in the rp 51 pre-mRNA is likely to act early in spliceosome assembly, and it was postulated that the intron enhancer may interact with U1 to facilitate early spliceosome assembly (Charpentier and Rosbash, 1996). Unfortunately, we did not observe any crosslinks between 4-thioU80 of U6 and the pre-mRNAs of these other yeast genes, in contrast to the crosslinks observed for actin pre-mRNA (data not shown). However, we did observe the usual crosslinking between 4-thioU 51 of U6 and the 5' splice sites of the other yeast gene pre-mRNAs. We cannot conclude that the intron-U6 crosslinks for 4-thioU80 is specific of an interaction found only in the actin gene. Perhaps, the introns of different yeast genes interact with U6 RNA in somewhat different conformations which which do not promote the crosslinking reaction using 4-thioU. Nonetheless, we have found the first evidence for an interaction between the conserved bulged nucleotide U80 in the 3' stem-loop of U6 and a nucleotide in the intron of the actin pre-mRNA which is part of an essential sequence for splicing in vitro. Uridine 80 in U6 RNA is a very highly conserved bulged nucleotide, and a phosphorothioate substitution 5' of this nucleotide or its analog results in loss of splicing in yeast, nematodes and in Group II splicing (Chanfreau and Jacquier, 1994; Fabrizio and

Abelson, 1992; Yu et al., 1995). While the conserved 3' stem-loop of U6 is known to be highly conserved and required for splicing, interactions of this helix with other spliceosomal components had not previously been identified.

It is interesting that a single 4-thioU placed at certain sites forms crosslinks to several RNAs. This suggests that the RNA-RNA interactions in the spliceosome may be dynamic in nature. This is especially true in light of the fact that 4-thioU is in effect a "zero-length" crosslinker, lacking any linker arms, and will only crosslink to RNAs in close proximity. Thus, it may be difficult to crosslink to two different RNAs simultaneously because of the distance contraint. In two cases (4-thioU51 and 80) we were able to fully characterize two different species crosslinked to a single nucleotide. 4-ThioU 51 in U6 crosslinks to both U4 snRNA and the 5' splice site of pre-mRNA/lariat intermediate, while 4-ThioU 80 crosslinks to both U4 snRNA and an essential element in the intron of the pre-mRNA. 4-ThioU does not necessarily form this many crosslinks at all positions in U6. In fact, when placed in some nucleotide positions, 4-thioU does not result in any crosslinks to RNA. This means that position 51 and 80 are sites where many critical interactions take place. Indeed, a base substitution of the invariant A51 to any other nucleotide causes a strong block of the second step of splicing (Fabrizio and Abelson, 1990), and the bulged U80 is a key conserved element of the U6 3' stem-loop (see discussion above). We also have a good approximation of the order of the U4/U6 crosslinks with respect to the pre-mRNA/U6 crosslinks. Since it is already known that U4 dissociates before the first step occurs (Yean and Lin, 1991), the (4-thioU51) U6/U4 snRNA crosslink presumably occurs early, prior to the first step splicing. This is very likely, since the U4/U6 crosslink we observed does not require the presence of exogenous pre-mRNA substrate. However, it may be more difficult to determine whether the U4/U6 (4-thioU51) crosslink occurs in U4/U6 snRNPs or U4/U6.U5 tri-snRNPs. We already know that the (4-thioU51) U6 RNA/ pre-mRNA crosslinks occur even in prp2 inactivated extracts, suggesting they occur prior to the first step of splicing in the A1 complex (Kim and Abelson, 1996). This

interaction is presumably maintained through the first step so that we find the 4-thioU51 crosslinked to the lariat intermediate after the first step. In the case of the 4-thioU80 crosslink, the crosslink to U4 snRNA probably occurs before the crosslink to pre-mRNA, since the U4/U6 crosslink was only ATP-dependent, but not dependent on pre-mRNA.

We have done an extensive search for crosslinks between U6 RNA and other snRNAs and pre-mRNA using 4-thioU site-specifically incorporated into many nucleotide positions in U6 RNA. We were able to do these experiments because we have an in vitro reconstitution system for U6 RNA. We obtained crosslinks between U6 and U4, which presumably form early in spliceosome assembly, and also crosslinks between U6 and U2 which are splicing reaction dependent. It is surprising that of the many positions that we tested, we only found one crosslink to an snRNA that was splicing-dependent. Either the incorporation of U6 RNA into spliceosomes is low, or the catalytic interactions of U6 RNA with other snRNAs are competed out by the much more efficient crosslinks to U4 RNA. We know that the incorporation of synthetic U6 RNA into spliceosomes is low since we find crosslinks between unincorporated U6 and RNAs in the extract -- these appeared in the absence of pre-mRNA and ATP. Moreover, the minimum amount of U6 needed for reconstitution is far greater than the amount of endogenous U6 RNA in the extract. In contrast to crosslinks between U6 RNA and other snRNAs, we obtained many more splicing-dependent crosslinks to the pre-mRNA. For the first time, we have shown that a conserved nucleotide in the 3' stem loop of U6 (an analogous to domain V in group II splicing), can crosslink to an important region in the pre-mRNA intron. The crosslinks we have obtained are highly sensitive markers for U6 RNA/snRNA interactions in splicing. Crosslinking with 4-thioU is a powerful biochemical technique because it allows us to capture interactions between U6 RNA and snRNAs or pre-mRNA that may be relevant to splicing. However, given the dynamic nature of the RNA interactions in the spliceosome, it may be difficult to model the RNA active site of the spliceosome using an observed crosslink as a constraint for the model in the context of previously determined interactions

(see Fig. 7b) since we do not know when some of these interactions occur. We may be able to determine the order of these crosslinks in spliceosome assembly by doing the experiments in prp mutant extracts blocked at well-defined stages in assembly or in the second step of splicing (Umen and Guthrie, 1995b). There are also deoxy substitution blocks to the first step of splicing which still allow assembly of the early spliceosomal complexes(Kim et al., 1997). We could also do the crosslinking experiments in glycerol gradient purified spliceosomes to determine the order of a particular 4-thioU crosslink (Yean and Lin, 1991). We can also determine if the interaction occurs when we knock out a protein that is thought to enhance a particular U6 RNA/snRNA interaction. Finally, we should be able to use these crosslinks as constraints in modeling the RNA active site of the spliceosome with a better knowledge of when they may be occurring in spliceosome assembly and splicing.

REFERENCES

Adams, C. J., Murray, J. B., Arnold, J. R. P., and Stockley, P. G. (1994). A convenient synthesis of S-cyanoethyl-protected 4-thiouridine and its incorporation into oligoribonucleotides. Tetrahedron Letters *35*, 765-768.

Adams, C. J., Murray, J. B., Farrow, M. A., Arnold, J. R. P., and Stockley, P. G. (1995). Incorporation of 6-thioguanosine into oligoribonucleotides. Tetrahedron Letters *36*, 5421-5424.

Ares, M., and Weiser, B. (1995). Rearrangement of snRNA structure during assembly and function of the spliceosome. Prog. Nucl. Acid Res. and Molec. Biol. *50*, 131-159.

Brow, D. A., and Guthrie, C. (1988). Spliceosomal RNA U6 is remarkably conserved from yeast to mammals. Nature (Lond.) *334*, 213-218.

Brow, D. A., and Vidaver, R. M. (1995). An element in human U6 RNA destabilizes the U4/U6 spliceosomal RNA complex. RNA *1*, 122-131.

Cavalier-Smith, T. (1991). Intron phylogeny: a new hypothesis. Trends Genet *7*, 145-148.

Cech, T. R. (1985). RNA splicing: three themes with variations. Cell *43*, 713-716.

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

Charpentier, B., and Rosbash, M. (1996). Intramolecular structure in yeast intron aids the early steps of in vitro spliceosome assembly. RNA *2*, 509-522.

Cortes, J. J., Sontheimer, E. J., Seiwert, S. D., and Steitz, J. A. (1993). Mutations in the conserved loop of human U5 snRNA generate use of novel cryptic 5' splice sites in-vivo. EMBO J *12*, 5181-5189.

Datta, B., and Weiner, A. M. (1991). Genetic evidence for base pairing between U2 and U6 snRNA in mammalian mRNA splicing. Nature (Lond.) *352*, 821-824.

Fabrizio, P., and Abelson, J. (1990). Two domains of yeast U6 small nuclear RNA required for both steps of nuclear precursor messenger RNA splicing. Science *250*, 404-409.

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

Fabrizio, P., McPheeters, D. S., and Abelson, J. (1989). *In vitro* assembly of yeast U6 snRNP: A functional assay. Genes Dev. *3*, 2137-2150.

Favre, A. (1990). 4-Thiouridine as an intrinsic photoaffinity probe of nucleic acid structure and interactions, H. Morrison, ed. (New York: John Wiley & Sons).

Fortner, D. M., Troy, R. G., and Brow, D. A. (1994). A stem/loop in U6 RNA defines a conformational switch required for pre-mRNA splicing. Genes Dev. *8*, 221-233.

Goguel, V., and Rosbash, M. (1993). Splice site choice and splicing efficiency are positively influenced by pre-mRNA intramolecular base pairing in yeast. Cell *72*, 893-901.

Hetzer, M., Wurzer, G., Schweyen, R. J., and Mueller, M. W. (1997). Trans-activation of group II intron splicing by nuclear U5 snRNA. Nature *386*, 417-420.

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

Kim, C. H., Ryan, D. E., Marciniec, T., and Abelson, J. (1997). Site-specific deoxynucleotide substitutions in yeast U6 snRNA block splicing of pre-mRNA in vitro. EMBO Journal *16*, 2119-2129.

Legrain, P., and Chanfreau, G. (1994). Pre-mRNA splicing: From intron recognition to catalysis. Bull. Inst. Pasteur *92*, 153-179.

Libri, D., Stutz, F., McCarthy, T., and Rosbash, M. (1995). RNA structural patterns and splicing: molecular basis for an RNA-based enhancer. RNA *1*, 425-436.

Lin, R. J., Newman, A. J., Cheng, S. C., and Abelson, J. (1985). Yeast pre-mRNA splicing in vitro. JBC *260*, 14780-14792.

Madhani, H. D., Bordonné, R., and Guthrie, C. (1990). Multiple roles for U6 snRNA in the splicing pathway. Genes Dev. *4*, 2264-2277.

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

Madhani, H. D., and Guthrie, C. (1994a). Dynamic RNA-RNA interactions in the spliceosome. Annu. Rev. Genet. *28*, 677-687.

Maschloff, K. L., and Padgett, R. A. (1992). Phosphorothioate substitution identifies phosphate groups important for pre-mRNA splicing. Nucleic Acids Res *20*, 1949-1957.

McSwiggen, J. A., and Cech, T. R. (1989). Stereochemistry of RNA cleavage by the Tetrahymena ribozyme and evidence that the chemical step is not rate-limiting. Science *244*, 679-83.

Moore, M. J., Query, C. C., and Sharp, P. A. (1993). Splicing of precursors to messenger RNAs by the spliceosome. In RNA World, R. Gesteland and J. Atkins, eds. (New York: Cold Spring Harbor Laboratory).

Moore, M. J., and Sharp, P. A. (1992). Site-specifically modified pre-mRNA: Substitutions of the 2'-hydroxyl groups at the splice sites. Science *256*, 992-997.

Moore, M. J., and Sharp, P. A. (1993). Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. Nature (Lond.) *365*, 364-368.

Newman, A. (1987). Specific accessory sequences in Saccharomyces cerevisiae introns control assembly of pre-mRNAs into spliceosomes. EMBO journal *6*, 3833-3839.

Newman, A. (1994). Small nuclear RNAs and pre-mRNA splicing. Curr Opin Cell Biol *6*.

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

Newman, A. J. (1997). The role of U5 snRNP in pre-mRNA splicing. EMBO J *16*, 5797-5800.

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

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

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

O'Keefe, R. T., Norman, C., and Newman, A. J. (1996). The invariant U5 snRNA loop 1 sequence isdispensable for the first catalytic step of splicing in yeast. Cell *86*, 679-689.

Parker, R., and Patterson, B. (1987). Architecture of fungal introns: Implications for spliceosome assembly, B. Dudock and M. Inouye, eds. (New York: Academic Press).

Parker, R., Siliciano, P. G., and Guthrie, C. (1987). Recognition of the TACTAAC box during mRNA splicing in yeast involves base-pairing to the U2-like snRNA. Cell *49*, 229-39.

Podar, M., Zhuo, J., Zhang, M., Franzen, J., Perlman, P. S., et al. (1998). Domain 5 binds near a highly conserved dinucleotide in the joiner linking domains 2 and 3 of a group II intron. RNA *4*, 151-166.

Rajagopal, J., Doudna, J. A., and Szostak, J. W. (1989). Stereochemical course of catalysis by the Tetrahymena ribozyme. Science *244*, 692-694.

Reich, C. I., Hoy, R. W. V., Porter, G. L., and Wise, J. A. (1992). Mutations at the 3' splice site can be suppressed by compensatory base changes in U1 snRNA in fission yeast. Cell *69*, 1159-1169.

Roiha, H., Shuster, E. O., Brow, D. A., and Guthrie, C. (1989). Small nuclear RNAs from budding yeasts: phylogenetic comparisons reveal extensive si ze variation. Gene *82*, 11269-11273.

Schumyatsky, G., and Reddy, R. (1993). Compilation of small RNA sequences. Nucleic Acids Research *21*, 3017.

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

Sharp, P. A. (1991). Five easy pieces. Science *42*, 663.

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.

Sontheimer, E. J., and Steitz, J. A. (1993). The U5 and U6 small nuclear RNAs as active site components of the spliceosome. Science *262*, 1989-1996.

Sontheimer, E. J., Sun, S., and Piccirilli, J. A. (1997). Metal ion catalysis during splicing of premessenger RNA. Nature *388*, 801-805.

Staley, J. P., and Guthrie, C. (1998). Mechanical Devices of the Spliceosome: Motors, Clocks, springs, and Things. Cell *92*, 315-326.

Tani, T., and Ohshima, Y. (1989). The gene for the U6 small nuclear RNA in fission yeast has an intron. Nature *337*, 87-90.

Umen, J. G., and Guthrie, C. (1995b). The second catalytic step of pre-mRNA splicing. RNA *1*, 869-885.

Vankan , P., McGuigan, C., and Mattaj, I. W. (1990). Domains of U4 and U6 snRNAs required for snRNAs required for snRNP assembly and splicing complementation in Xenopus oocytes. EMBO J *9*, 3397-3404.

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

Wise, J. A. (1993). Guides to the heart of the spliceosme. Science *262*, 1978-1979.

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

Wu, J., and Manley, J. L. (1991). Base pairing between U2 and U6 snRNAs is necessary for splicing of a mammalian pre-mRNA. Nature (Lond.) *352*, 818-821.

Yean, S. L., and Lin, R. J. (1991). U4 small nuclear RNA dissociated from a yeast spliceosome and does not participate in the subsequent splicing reaction. Mol. Cell. Biol. *11*, 5571-5577.

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

Yu, Y. T., Maroney, P., and Nilsen, T. W. (1993). Functional reconstitution of U6 snRNA in nematode cis and trans-splicing: U6 can serve as both a branch acceptor and a 5' exon. Cell *75*, 1049-1059.

Zhuang, Y., and Weiner, A. M. (1986). A compensatory base change in U1 snRNA suppresses a 5' splice site mutation. Cell *46*, 827-35.

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.