The Role of U1 snRNP in Pre-mRNA Splicing
by
John D. Crispino

Submitted to the Department of Biology on November 28, 1995
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

ABSTRACT

The role of U1 snRNP in splicing is twofold: first, U1 snRNP recognizes the 5' splice site by base pairing and facilitates subsequent interactions of these sequences with other factors, including U6 snRNA; second, U1 snRNP promotes the binding of U2 snRNP to the branch site in the formation of the A complex. This thesis provides evidence that high concentrations of SR proteins circumvent the requirement for U1 snRNP in the pre-mRNA splicing reaction. That SR proteins permit processing of pre-mRNA in reactions depleted of U1 snRNP suggests that U1 snRNA is not essential in splicing for catalysis, 5' splice site selection, or spliceosome assembly.

SR proteins promote efficient splicing in extracts that have been depleted for U1 snRNP. This bypass reaction is substrate specific, probably reflecting the efficiency of binding of SR proteins. In this reaction, SR proteins first promote the rapid association of U2 snRNP with the branch site, in a manner independent of U1 snRNP or the 5' splice site. The entire 5' splice site consensus sequence is required in a step after the formation of complex A. This recognition of the 5' splice site occurs in a rate-limiting reaction with or after the binding of the U4/U5/U6 tri-snRNP. SR proteins may have a role in recruiting the tri-snRNP as well.

Two precursor RNAs are active for splicing in U1 snRNP-depleted extracts without a requirement for additional SR proteins. These substrates may contain high affinity binding sites for SR proteins, or alternatively, the splicing of these RNAs may represent a different U1 snRNP bypass reaction. Several cis-acting sequences that are necessary for U1-independent splicing have been identified in these substrates. The characterization of factors that recognize these sequences should lead to a better understanding of the U1 bypass mechanism.

In vivo, most pre-mRNAs are processed in the U1-dependent manner, but there may exist alternative pathways for splicing. U1 snRNP could have a limited role in the splicing of some pre-mRNAs, specifically those that have a high affinity for SR proteins. This secondary splicing pathway would give the cell an additional way to regulate the splicing of different classes of pre-mRNAs.

Thesis Supervisor: Dr. Phillip A. Sharp, Salvador E. Luria Professor and Head, Department of Biology
# TABLE OF CONTENTS

Abstract........................................................................................................... 2
Dedication........................................................................................................... 4
Acknowledgments.............................................................................................. 5

Chapter I

Introduction and overview.............................................................................. 6
U1 snRNP Functions in splicing................................................................. 9
U1-independent splicing.............................................................................. 14
The SR protein superfamily........................................................................ 16
  mAb104 family...................................................................................... 17
  ASF/SF2............................................................................................... 20
  SC35....................................................................................................... 21
  SR related proteins.............................................................................. 22
  U2AF and U1 70K Splicing factors.................................................... 22
Mechanisms of alternative splicing.............................................................. 23
  Drosophila sex determination.......................................................... 23
  Exon definition.................................................................................... 27
  Exon enhancers.................................................................................. 31
    Bovine FP......................................................................................... 32
    Fibronectin ED1............................................................................... 33
    Cardiac Troponin T.......................................................................... 33
    Mouse IgM....................................................................................... 34
  5' splice site selection.......................................................................... 35
References....................................................................................................... 37

Chapter II

Complementation of U1 snRNP depleted extracts by SR proteins.............. 43

Chapter III

A pre-mRNA:U6 snRNA interaction can be rate limiting for U1-independent splicing........................................................................ 62

Chapter IV

U1-independent splicing in the absence of excess SR proteins.................. 100

Appendix

Mapping the sequences important for SR complementation of U1 snRNP-depleted extracts............................................................... 125

Biographical Note.......................................................................................... 146

3
To Ellen
I am very grateful to Phillip Sharp for providing me with the opportunity to work in his laboratory. Phil has been an ideal advisor for me: never too intrusive, but never too busy to discuss results and provide direction for my research. I thank him for his continuous support and advice.

The laboratory wouldn’t be such a sane place without Margarita Siafaca. I will miss our daily discussions about cars, coffee, work, houses, “pets” and life in the suburbs. I appreciate her friendship and look forward to future wine and coffee meetings.

I wish to also thank the members of my thesis committee. Professor Tony Sinskey has been a great ally and teacher. I am grateful for all his advice on the pharmaceutical and biotech industries. I also extend my thanks to Professor Tom RajBhandary for sharing his enthusiasm about pre-mRNA splicing and biology in general. It was always a pleasure to meet with him to discuss both science and my career goals. Additional thanks to Professor Frank Solomon for joining my thesis committee and reading my thesis. Finally, I am glad that Professor Michael Green of the University of Massachusetts agreed to participate in my thesis defense, and I am looking forward to spending several months working with Dr. Green at Scriptgen.

Special thanks to Ben Blencowe for sharing his expertise and ideas. His arrival in the Sharp lab couldn’t have been timed any better. Thanks for teaching me how to deplete nuclear extracts and being such a great collaborator.

I wouldn’t have felt so comfortable in the lab so quickly without the help of Melissa Moore. Her enthusiasm pulled me into the Sharp lab and motivated me to work hard. I’m sorry that our project didn’t work out, but I would do it all over again. Thanks for getting me started on the right tract.

Rachel Meyers also helped me feel at home in our bay within the Sharp lab. Her advice and friendship was invaluable. I look forward to working with her in the future. I also am grateful to Charles Query, my bench mate for the first few months. I thank him for offering assistance and sharing his enthusiasm for research.

I’m glad that I have gotten to know Joel “Royale with cheese” Pomerantz and Andrew “Bud the Spud” MacMillan so well in my last year of graduate school. Both helped me to maintain my sanity while writing post-doctoral fellowships and this thesis. Thanks for the movies, the coffee, the burgers and the music.

I also want to thank the other members of the Sharp lab who have helped make graduate school worthwhile, including my fellow graduate students, Ben Shykind, Jae Kim, Patrick McCaw, Sandy Gilbert, and Lee Lim. I look forward to being colleagues again someday. It was a pleasure working with the intelligent post-docs that occupied space in the Sharp lab, such as Sean “I can catch fish, although never when you’re with me” Harper, Dan “fellow Ward Melvillite” Chasman, and bay-mate Qiang Zhou, as well as Jorgen Kjems, Jeff Parvin, Anna Gil, Tom Kristie, Marc Timmers, Ken LeClair, David Fisher, David Potter and Marcia Morton. Best wishes to those post-docs who have recently arrived that will soon be famous, Karyn Cepek, Rock Pulak, Matt Smith and Tom Tuschl.

I am very grateful to Robbyn Issner for being such a good friend. I will miss our early morning coffee breaks at ABP. Similarly, I thank Yubin Qiu for her intelligence and bright outlook. It was a pleasure working with them.

Finally, I wish to thank my wife and my family for their constant support of my graduate education.
CHAPTER I

INTRODUCTION AND OVERVIEW
Eukaryotic genes are frequently interrupted by intervening sequences, termed introns, which must be removed before protein synthesis occurs. The precise excision of these sequences is necessary for gene expression, because errors in splicing produce truncated or mutant proteins with altered activities. Introns are deleted from the precursor RNA by two transesterification reactions (Figure 1A). In the first step, a conserved adenosine residue within the intron, termed the branch site, is joined to the 5' end of the intron, producing the lariat intermediate and releasing the 5' exon. The second step consists of cleavage at the 3' splice site with concomitant joining of the exons. This removal of introns from nuclear pre-mRNAs occurs in a large multi-subunit complex termed the spliceosome. The spliceosome, which is composed of four snRNP particles (U1, U2, U4/U6 and U5) and many non-snRNP protein factors, forms in a stepwise manner on the substrate (Figure 1B; for review, see Moore et al., 1993; Madhani and Guthrie, 1994). In the first commitment step, U1 snRNP, in conjunction with the splicing factor U2AF and several members of the SR protein family, associates with the pre-mRNA to form the E complex (Michaud and Reed, 1993; Staknis and Reed, 1994). Next, U2 snRNP binds to the branch site, forming the A complex (Konarska and Sharp, 1986; 1987). Finally, the assembly of the spliceosome is completed by the addition of the U4/U5/U6 tri-snRNP.

The question of how splice sites are recognized has been a topic of intense research. During splicing, the 5' splice site is first recognized by the U1 snRNP, which binds to the substrate by base pairing (Zhuang and Weiner, 1986; Seraphin et al., 1988; Siliciano and Guthrie, 1988). The binding of the U1 snRNP to this sequence does not sufficiently define the splice site, as U1 snRNP often binds equivalently to competing 5' splice sites (Eperon et al., 1993; Tarn and Steitz, 1994; Zahler and Roth, 1995). Additional protein factors, including members of the SR family, are probably also involved in splice site selection.
This thesis describes work performed to investigate the role of the U1 snRNP in splicing. U1 snRNP-depleted nuclear extracts were made by the antisense affinity selection method (Blencowe and Barabino, 1995). These extracts were not active for the splicing of added pre-mRNA and proved to be valuable reagents for the study of U1 snRNP. In the course of attempting reconstitution of splicing activity by U1 snRNA and various protein fractions derived from HeLa cell extracts, it was observed that a highly purified preparation of SR proteins fully restored splicing activity to these snRNP-depleted extracts (Crispino et al., 1994). This thesis describes the SR-mediated splicing of U1 snRNP-depleted extracts and provides strong evidence that this bypass reaction proceeds in a manner independent of this snRNP.

Chapter 2 details the reconstitution of U1 snRNP-depleted extracts by high concentrations of SR proteins. The addition of a highly purified preparation of SR proteins fully restored splicing activity in reactions depleted of U1 snRNP in a substrate specific manner. Affinity selection experiments demonstrated that spliceosomes lacking U1 snRNA formed in these reactions (Crispino et al., 1994).

Chapter 3 addresses the mechanism by which splicing occurs in the absence of U1 snRNP. In this bypass reaction, the SR proteins first promote the rapid association of U2 snRNP with the branch site, in a manner independent of either U1 snRNP or the 5' splice site. The recognition of the 5' splice site then occurs in a rate-limiting step with or after the binding of the U4/U5/U6 tri-snRNP (Crispino and Sharp, 1995).

Chapter 4 focuses on understanding the sequence specificity of the U1 snRNP bypass reaction. Splicing reactions with the Ftz pre-mRNA from Drosophila as well with as a synthetic substrate, PIPβG-E', are summarized in this chapter: these results demonstrate that splicing can occur in U1 snRNP-depleted extracts even with the
endogenous level of SR proteins. This splicing pathway provides a means by which U1-independent splicing could occur under physiologically relevant concentrations of SR proteins.

Finally, the appendix summarizes results from experiments aimed at characterizing the sequences that promote the SR-mediated reconstitution of U1 snRNP-depleted extracts. Surprisingly, the regions that confer the ability to splice in the absence of U1 snRNP with SR proteins are located in the intron and neither the exons, nor the 5' splice site.

The intention of this chapter is to provide a context for these results, especially with respect to constitutive and alternative mechanisms of 5' splice site selection. The sections focus on 1) the functions of U1 snRNP in splicing; 2) the role of SR proteins as constitutive and alternative splicing factors; 3) the mechanisms of alternative splicing; 4) 5' splice site selection.

U1 snRNP

Functions in Splicing

The original proposal that U1 snRNP is involved in pre-mRNA splicing was based on the base-pairing potential between the 5' end of U1 snRNA and pre-mRNA splice junctions (Lerner et al., 1980; Rogers and Wall, 1980). Experimental evidence that U1 snRNP is involved in splicing includes a wealth of both biochemical and genetic observations. First, anti-U1 snRNP antibodies inhibited splicing both in vivo (in Xenopus oocytes) and in vitro. In addition, antibody mediated depletion of the snRNP, although not absolutely specific for the U1 particle, rendered the extracts inactive for splicing. Second, isolated U1 snRNPs were detected binding to the 5' splice site of pre-mRNA substrates in vitro, thereby protecting these regions from RNase digestion. Third,
Figure 1A. Pre-mRNA splicing reactions. Pre-mRNA splicing occurs via a two-step transesterification reaction. The first reaction results in cleavage at the 5' splice site to yield a 5' exon intermediate and the intron with attached 3' exon. This second intermediate is a branched RNA in which a conserved adenosine within the intron is covalently linked by a 2'-5' phosphodiester bond to the first nucleotide of the intron. In the second transesterification, the free 3' OH group of the 5' exon becomes ligated to the second exon at the 3' splice site to produce messenger RNA with concomitant release of the lariat product.

B. Spliceosome assembly. The two splicing reactions occur within the confines of a large multi-subunit complex called the spliceosome, which forms in a step-wise manner on the substrate. In the earliest stage of assembly, the splicing factors U2AF, members of the SR family, and U1 snRNP bind to the pre-mRNA to form the commitment complex (CC; E complex). In the first ATP dependent step, U2 snRNP associates with E complex by base-pairing with the branch site to form the A complex. U1 snRNP and SR proteins remain bound at this stage, whereas U2AF is displaced from the substrate. Finally, the U4/U5/U6 tri-snRNP enters the complex to form the B/C spliceosome. The splicing reactions proceed only after the formation of the complete spliceosome.
A

Pre-mRNA

Lariat Intermediate

mRNA Product

5' GpG 3'

Lariat Product

UACUAC AG-OH 3'
pre-mRNA

+ U1 snRNP

E complex

+ ATP
+ U2 snRNP

A complex

+ ATP
+ U4/5/6 snRNP

B/C complex
selective degradation of the 5' end of U1 snRNP, the sequence which recognizes the pre-mRNA, by RNase H cleavage abolished splicing (for review, see Steitz et al., 1988).

More recently, antisense RNA oligonucleotides have been used to block the 5' end of U1 snRNP, rendering the extracts unable to splice exogenously added substrates. Nuclear extracts have also been depleted of U1 snRNP by affinity chromatography using biotinylated antisense oligonucleotides; the depletion of the snRNP from nuclear extracts also inhibited splicing (Barabino et al., 1990).

Several genetic studies have demonstrated that the 5' end of U1 snRNA associates with the 5' splice site by base pairing (Zhuang and Weiner, 1986; Siliciano and Guthrie, 1988). Some mutations of the 5' splice site were suppressed by compensatory changes in U1 snRNA, but in many cases the compensatory changes did not restore splicing activity (Seraphin et al., 1988; Siliciano and Guthrie, 1988; Seraphin and Rosbash, 1990).

U1 snRNP recognizes the 5' splice site early in spliceosome assembly. Although it is not stably bound to the substrate under non-denaturing gel electrophoresis (Konarska and Sharp, 1986; 1987), the snRNP has been detected in a complex with pre-mRNA substrates by affinity purification (Bindereif and Green, 1987). Gel filtration chromatography has also allowed for the purification from mammalian extracts of a complex which contains U1 snRNP, U2AF, and SR proteins (E complex; Staknis and Reed, 1994). A similar ATP-independent commitment complex containing U1 snRNP has been defined in yeast (CC complex; Seraphin and Rosbash, 1989; 1991).

Subsequent to the formation of the commitment complex, U1 snRNP promotes the stable binding of the U2 snRNP to the branch site to produce complex A, while remaining associated with the pre-mRNA (Michaud and Reed, 1991). The 5' end of U1 snRNA is not required for the U1 snRNP mediated stimulation of U2 snRNP binding, as extracts in which the 5' end was blocked by an antisense oligo supported complex formation (Barabino et al., 1990).
The complete spliceosome is formed by the addition of the U4/U5/U6 tri-snRNP to complex A. At this stage, the U1 snRNA is probably displaced from the 5' splice site (Konforti et al., 1993) to allow for the binding of U6 snRNA. Recent experiments have revealed an interaction between U6 snRNA and distal nucleotides (+4 through +6) of the 5' splice site by detection of allele specific complementation between mutations in the 5' splice site and mutations in U6 snRNA and also by ultraviolet light-induced cross-links between these two RNAs (Lesser and Guthrie, 1993; Kandels-Lewis and Seraphin, 1993; Sontheimer and Steitz, 1993; Sun and Manley, 1995).

The distal nucleotides in the 5' splice site are therefore recognized at least twice during the splicing reaction, by U1 and U6 snRNAs. The other nucleotides in the 5' splice site may also be recognized a second time, probably in the binding of the U4/U5/U6 tri-snRNP. The observation that these 5' splice site sequences are recognized subsequent to U1 snRNA recognition explains the previous lack of complementation by changes which restore U1 base pairing. These results are consistent with the hypothesis that U1 snRNP's primary role is to promote the assembly of U2 snRNP on the branch site, perhaps by facilitating the binding of SR proteins to the substrate RNA.

U1-independent splicing

The data presented in this thesis give strong evidence that splicing reactions can proceed in the absence of U1 snRNP under certain conditions. This finding was anticipated by previous studies that show that U1 snRNP is easily displaced from the spliceosome (Konarska and Sharp, 1986; 1987) and that there is a lack of an obvious analog to U1 snRNA in trans-splicing reactions (Nilsen, 1993). In fact, early studies with Xenopus oocytes demonstrated that an antibody against U1 snRNP was unable to block the splicing of some 5' splice sites (Fradin et al., 1984). Further, splicing of the SV40 large T intron and chimeric RNAs with sequences from adenovirus substrate fused to
spliced leader sequences were shown to be less sensitive to RNase H degradation of U1 snRNA (Bruzik and Steitz, 1990).

This thesis demonstrates that SR proteins promote efficient splicing in extracts that have been depleted for U1 snRNP. In a separate study, SR proteins were shown to reconstitute the splicing of pre-mRNAs in extracts where the U1 snRNP was blocked by an antisense oligonucleotide. In these reactions, no U1 snRNA was detected interacting with the 5' splice site by psoralen cross-linking (Tarn and Steitz, 1994).

That high concentrations of SR proteins permit processing of pre-mRNA in reactions depleted of U1 snRNP suggests that U1 snRNA is not essential for catalysis of either of the transesterification steps in splicing. Studies with yeast extracts support this model: U1 snRNA functioned only to assemble the spliceosome and was dispensable for the subsequent catalysis (S.-L. Yean and R.-J. Lin, personal communication). Our data suggest that even this role of U1 snRNP can be mediated by SR proteins. Further evidence that the spliceosome can form without U1 snRNP is provided by studies showing that spliceosome-type complexes containing U2, U5, and U4/U6 assemble on short oligoribonucleotides encompassing a 5' splice site (Konforti et al., 1993). Formation of these complexes was dependent on the recognition of the 5' splice site but did not depend on the presence of an intact U1 snRNA.

Recent studies of trans-splicing have suggested that U1 snRNP is not required and actually inhibits the joining of exons from two different RNAs (Konforti and Konarska, 1995). Splicing of RNA substrates in a trans-fashion has been observed in several systems, including HeLa cells, nematodes, and trypanosomes (for review, see Moore et al., 1993). Initial experiments suggested that the two splice site RNAs interact to promote the splicing process (Konarska et al., 1985; Solnick, 1985). More recently, trans-splicing was shown to be stimulated by SR proteins (Bruzik and Maniatis, 1995) and dependent on splicing enhancers (see below; Chiara and Reed, 1995). The role of U1 snRNP in these cases has not yet been investigated. In contrast, experiments with
minimal substrates have demonstrated that U1 snRNP is not required for the reaction. The short 5' splice site oligonucleotides that form spliceosome type complexes undergo both steps of splicing when a second RNA containing the 3' splice site region is added in trans (Konforti and Konarska, 1995). Interestingly, the trans-splicing reaction between the short 5' splice site oligonucleotides and a 3' splice site RNA is enhanced in HeLa nuclear extracts when the 5' end of U1 snRNA is blocked or when the U1 snRNP is removed. Thus, the requirement for U1 snRNP is bypassed in this trans-splicing reaction. In addition, these U1-independent reactions do not require elevated levels of SR proteins.

The primary role of U1 snRNP, therefore, is to promote the binding of U2 snRNP to the branch site. U1 snRNP bound at either an upstream 5' splice site, or at a downstream 5' splice site is likely to commit a pre-mRNA to splicing and to stimulate the binding of U2 snRNP. In the absence of U1 snRNP, SR proteins can substitute by facilitating the binding of U2 snRNP, possibly through the association with a splicing factor called U2AF. These models of spliceosome assembly will be described in greater detail in the following section.

THE SR SUPERFAMILY OF SPlicing FACTORS

The recent identification of a group of splicing factors that share a common motif of alternating serine and arginine residues has provided new insight into the mechanism of both constitutive and alternative splicing. The members of this extended family of splicing factors, the SR proteins, are important components of the splicing machinery in nearly every stage of spliceosome assembly. These proteins function constitutively in splicing to promote the recognition of pre-mRNAs and to facilitate the formation of the commitment complex (E), the A complex, and the mature spliceosome including the U4/U5/U6 tri-snRNP (B/C complexes; Kohtz et al., 1994; Staknis and Reed, 1994; Crispino and Sharp, 1995; Roscigno and Garcia-Blanco, 1995; Tarn and Steitz, 1995).
addition, these proteins play a role in the regulation of alternative splicing, with the best example being the control of sex determination in *Drosophila*. This section will describe important members of the family and summarize the known and putative functions of these proteins in splicing.

**The mAbl04 Family**

While the SR family was originally described by three *Drosophila* splicing regulators, *suppressor of white apricot* (SWAP), *transformer* (tra) and *transformer-2* (tra-2), the best characterized members of this family belong to a group of proteins which share a common epitope recognized by an antibody, mAbl04, that stains lampbrush chromosomes (Roth et al., 1990). Six mAbl04 reactive proteins co-purify in a two-step procedure from nuclear extract which consists of a 65%-90% ammonium sulfate selection followed by precipitation in 20 mM MgCl₂ (Zahler et al., 1992). These proteins share a common bipartite structure consisting of an RNA binding domain (RRM; RNA recognition motif; Query et al., 1989) and the SR domain (Figure 2). In addition, these proteins are conserved from *Drosophila* to humans in both size and structure. An important criteria of these SR proteins is their ability to reconstitute the splicing of the S100 cytoplasmic fraction, which contains all the snRNPs required for splicing, but is blocked at the earliest stage of spliceosome assembly. Each of the six mAbl04 proteins can complement S100 splicing, indicating a degree of overlapping function.

Several of these proteins have been extensively characterized, with the best examples being SRp30a, also known as ASF/SF2, and SRp30b, identified as SC35. The following sections describe these proteins in greater detail; the discussion then turns to characterizing the extended family of SR proteins and the additional roles that these proteins play in the regulation of gene expression.
Figure 2. Schematic of SR protein structure. The most well defined members of the SR protein family, those which cross-react with the mAb104 antibody, consist of two domains. Each protein contains a consensus RNA binding domain, the RRM (the RNA recognition motif). Several of these proteins also have a second RNA binding domain that is homologous to the RRM. In addition, SR proteins have an extended region of alternating serine and arginine residues which is most likely involved in protein-protein interactions.
The SR104 Protein Family

RRM  RRM-H  Ser-Arg Motif
SRp75
SRp55
SRp40
SRp30a (ASF/SF2)
SRp30b (SC35)
SRp20
ASF/SF2

SRp30a is identical to ASF and SF2 which were independently isolated as splicing factors. ASF was identified as an activity from 293 cells that altered the usage of competing 5' splice sites in the SV40 early pre-mRNA; the addition of this factor to HeLa extracts promoted splicing of the small t 5' splice site (the proximal site; Ge and Manley, 1990). In a separate study, the 33kD factor SF2 was purified from HeLa cells as a protein which could complement an S100 extract (Krainer et al., 1990). SF2 was also found to influence 5' splice site selection; high concentrations of SF2 promoted the use of the 5' splice site closest to the 3' splice site. The two activities were cloned and found to be identical to one another, showing homology to the Drosophila proteins SWAP, Tra and Tra-2 (Ge et al., 1991; Krainer et al., 1991).

The alternative splicing activity of ASF/SF2 is antagonized by the hnRNP A1 protein. In vitro, excess concentrations of hnRNP A1 drive the splicing of the SV40 pre-mRNA to the distal large T 5' splice site (Mayeda and Krainer, 1992). Thus, in vitro the relative concentrations of these two proteins precisely determine the 5' splice site. Interestingly, this situation is recapitulated in vivo. Increased expression of ASF/SF2 in HeLa cells activated proximal 5' splice sites, while the overexpression of hnRNP A1 had the opposite effect on a transfected reporter (Caceres et al., 1994). The ASF/SF2 effect was observed on several substrates, including the thalassemic allele of β-globin and the rat clathrin light chain B, which shows neuron specific expression. These results suggest that changes in the levels of splicing regulators, such as ASF/SF2, may control the expression of many classes of genes.

The two activities of ASF/SF2 are distinct in their sequence requirements. Mutations in the RS domain disrupt the constitutive splicing activity, but do not alter the splice site switching activity (Caceres and Krainer, 1993; Zuo and Manley, 1993). Thus, the binding of ASF/SF2 to the pre-mRNA appears to be the critical step in its promotion of proximal splice site usage. In contrast, the reconstitution of S100 requires both an
intact SR domain and RNA binding motif. Therefore, distinct molecular interactions are involved in the two activities.

ASF/SF2 may also participate in 5' splice site selection by promoting the binding of U1 snRNP to the 5' splice site. This cooperation probably occurs by specific interactions between its SR domain and a similar region in the U1 snRNP specific protein U1 70K (Kohtz et al., 1994). ASF/SF2 was initially reported to specifically bind to sequences in the 5' splice site (Zuo and Manley, 1994), although most 5' splice sites were not targets of high affinity binding (Tacke and Manley, 1995). More recently, the RNA binding specificity was determined using in vitro selection (Tacke and Manley, 1995). The optimized binding sites were purine rich, explaining the observation by Zuo and Manley whose 5' splice sites were purine rich.

SC35

A monoclonal antibody raised against mammalian spliceosomes specifically recognized a 35kD doublet in Western blots of nuclear extracts. The antigen, SC35, is required for the first step in the splicing reaction and for spliceosome assembly, as splicing in extracts treated with anti-SC35 is inhibited at early steps (Fu and Maniatis, 1990). The domain structure is similar to that of the other mAb104 reactive proteins (Figure 2). Although the two proteins display only 31% amino acid identity, SC35 has a similar effect on competing 5' splice sites as ASF/SF2, and can also be antagonized by hnRNP A1. SC35 also efficiently promotes splicing in the deficient S100 cytoplasmic extract.

SC35 functions during an early step in the commitment of a substrate to splicing (Fu, 1993), in the formation of the E complex (Staknis and Reed, 1994), and in bridging interactions between the 5' splice site and the branch site (Wu and Maniatis, 1993). Figure 1 illustrates the current model of the commitment complex; it is competent to proceed to complex A upon the binding of U2 snRNP. SC35 appears to play a role in
recruitment of U2 snRNP even in the absence of U1 snRNP, perhaps by facilitating the interaction of U2AF with the branch site or by direct recruitment of U2 snRNP (Crispino and Sharp, 1995; Tarn and Steitz, 1995).

Purified SC35 was shown to bind RNA non-specifically (Fu and Maniatis, 1992). A more detailed analysis of its RNA binding specificity by in vitro selection showed that although the preferred site was purine rich, it was distinguishable from that of ASF/SF2 (Tacke and Manley, 1995).

Staining of the nucleus with the anti-SC35 antibody revealed that the protein is concentrated in discrete regions of the nucleus, called speckles (Fu and Maniatis, 1990; for review, see Spector, 1993). Each of the splicing snRNPs, as well as other members of the SR family, are also localized in speckles. Whether these subnuclear regions are sites of splicing or merely a storage facility for splicing factors is currently a matter of debate (for review, see Fu, 1995).

SR Related Proteins

Recent studies utilizing monoclonal antibodies raised against the human nuclear matrix revealed several high molecular weight antigens which have properties similar to SR proteins (Blencowe et al., 1994; 1995). One of these antigens, B1C8, has been cloned and found to contain an SR domain (Blencowe and Sharp, personal communication). Similarly, a monoclonal antibody raised against four of the defined SR proteins recognized over 20 distinct proteins in nuclear extracts, with many of these factors detected in spliceosomes (Neugebauer et al., 1995). Many of these SR related proteins are likely to be important factors in RNA processing.

U2AF and U1 70K Splicing Factors

Two well defined splicing factors contain domains composed of alternating arginine serine residues, although they do not co-purify with the mAb104 subset, nor do
they cross react with the antibody. U1 70K is an integral component of the U1 snRNP that contains a prototypic RRM and a region that is rich in arginine and serine residues (Query et al., 1989). As mentioned above, the 70K factor interacts with the SR proteins ASF/SF2 (biochemical data) and SC35 (yeast two hybrid assay). While the U1 70K protein is not thought to bind the pre-mRNA on its own, it is tethered to the 5' splice site by virtue of being associated with U1 snRNA. Thus, it is positioned to allow for communication with the 3' splice site, probably through SC35.

The U2 auxiliary factor (U2AF65) binds the polypyrimidine tract independent of other splicing factors, and its binding is required for the subsequent association of U2 snRNP with the branch site (Zamore et al., 1992). The RNA binding region is composed of three ribonucleoprotein consensus motifs which direct the specific binding to U-rich RNAs (Singh et al., 1995). In addition, a second subunit of U2AF, U2AF35, is involved in spliceosome formation, probably by associating with other SR proteins including the SC35 factor (Wu and Maniatis, 1993).

The relative affinities of U2AF65 for different polypyrimidine tracts may in part explain differences in the efficiency of splicing of cellular RNAs. RNAs containing weak polypyrimidine tracts don't bind U2AF as well, providing the cell with another means to regulate splicing. Examples of this mechanism of splicing regulation are described in the next section.

MECHANISMS OF ALTERNATIVE SPLICING

*Drosophila* Sex Determination

Perhaps the best characterized example of alternative splicing is the process of sex determination in *Drosophila* (Figure 3). In this pathway, a cascade of regulated splicing events directs the development of the fly as either male or female. The primary determinant of sex is the ratio of X chromosomes to autosomes. A female ratio (2X:2A)
leads to the activation of transcription of the *Sex lethal* (*Sxl*) gene. The Sxl protein, a splicing factor that contains an RNA recognition motif but lacks an SR domain influences the splicing of two genes: its own and that of the *transformer* (*tra*) locus. First, Sxl dictates the splicing of its own transcripts by promoting the exclusion of exon 3; this maintains the female specific protein. In males, the default pattern is the inclusion of this exon, which contains a stop codon and thus synthesizes a truncated non-functional protein. Second, Sxl regulates the alternative splicing of the Tra pre-mRNA by directing the exclusion of exon 2 in females. Once again, the default splicing pattern includes an exon which contains a stop codon that produces a non-functional protein in males.

The Sex lethal protein acts as a negative regulator of splicing by antagonizing the activity of the essential factor U2AF (Valcarcel et al., 1993). Sxl binds to sequences upstream of the 3' splice site of exon 2, thereby blocking the association of U2AF with the polypyrimidine tract. This enables U2AF to activate the female specific site that has a lower affinity for U2AF. Studies of the RNA binding of Sxl have demonstrated that the consensus is similar to that of U2AF in its uridine richness, but that it is a distinct element. The majority of selected sequences were composed of a stretch of 17 to 20 uridines interrupted by two to four guanosines (Singh et al., 1995). Interestingly, a chimeric protein composed of the Sxl RRMs fused to the SR domain of U2AF promoted splicing to exon 3, instead of blocking it (Valcarcel et al., 1993). This suggests that the absence of an effector domain on Sxl confers a dominant negative activity on the protein. It is also demonstrates that these splicing factors are modular, in the same way that the majority of transcription factors are composed of discrete DNA binding and effector domains. Splicing factors may therefore be amenable to manipulation for the production of chimeric proteins which could regulate the splicing of specific genes (Pomerantz, 1995).

In the next step of the cascade, female-specific Transformer acts to positively affect the splicing of the *doublesex* (*dsx*) gene. The Tra protein, in concert with another
**Figure 3.** *Drosophila* sex determination pathway. A cascade of regulated splicing events is responsible for the expression of sex-specific proteins that dictate the sex of the mature fly. The female Sex lethal (Sxl) protein represses inclusion of the second exon in the pre-mRNA of the *transformer* (*tra*) gene to produce the female-specific Tra protein. This protein, in turn, activates the inclusion of exon 4 in *doublesex* (*Dsx*) mRNA which produces the female-specific Dsx protein. The details of splicing regulation are described in the text. This figure was adapted from Moore et al., 1993.
protein, Tra-2, activates the 3' splice site of the Dsx exon 4, which is normally excluded from messenger RNA. Male and female specific Dsx proteins share the first three exons, but have different 3' exons and polyadenylation sites. These proteins are postulated to be transcriptional repressors; female Dsx represses male differentiation, and male Dsx represses female differentiation (Burtis et al., 1991).

The alternative splicing of \textit{dsx} requires an RNA splicing enhancer, designated the \textit{dsx} repeat element, comprised of six copies of a 13-nt repeat sequence (Hedley and Maniatis, 1991). These repeats are sites for binding of the Tra-2 protein, which binds via its RRM; the Tra protein lacks an RNA binding domain and is thought to associate with the Tra-2 via protein-protein interactions (Tian and Maniatis, 1993). Yeast two hybrid experiments have confirmed this model and in addition have suggested that both Tra and Tra-2 can associate with other SR proteins, presumably through their SR motifs (Wu and Maniatis, 1993). More recently, a purine rich element within the \textit{dsx} repeat element has been identified as being required for proper splicing regulation (Lynch and Maniatis, 1995). The Tra proteins and additional SR factors bind cooperatively to this element. The mechanism of splicing activation will be described in subsequent sections which discuss splicing enhancers. These purine rich sequences are thought to be splicing enhancers, in some ways analogous to transcription enhancers in that they can promote the splicing process. The next section focuses on both the sequences and the putative mechanism of their action.

\textbf{Exon Definition}

The removal of an intron from a pre-messenger RNA depends on the integrity of sequences at both the 5' and 3' ends of the intron, as mutations at either site impair excision of the intron (Lamond et al., 1987). These sequences are necessary, in part, for the efficient binding of splicing factors that, in turn, interact with one another to dictate the precise location of the splicing reactions. Such collaboration between splice sites has
been observed in many in vitro studies where factors bound at the 5' splice site interacted with those bound at the downstream 3' splice site (for example, E complex; Figure 1). The association between U1 snRNP bound at the 5' splice site and U2AF at the downstream 3' splice site is known as intron definition. In this model, the intron is the unit of recognition; the spliceosome assembles at both ends of the intron with factors communicating across the intron (Figure 4A). An alternative model, exon definition, proposes that exons are the unit of recognition (Berget, 1995). Here, factors bound at a 3' splice site are proposed to interact with others bound at a downstream 5' splice site (Figure 4B). In this latter model, the associations between splicing factors span the exon sequences.

In simple two exon substrates with short introns, the first model is sufficient to describe the splicing patterns. Intron definition, however, does not adequately explain the size constraint of mammalian exons, or observations made from experiments utilizing multiple-intron substrates (for review see Moore et al., 1993; Berget, 1995). In vertebrates, exons are usually short (averaging 137 nucleotides) while introns are large, with some approaching 100,000 nts. This small exon size appears to be important for efficient splicing, because introns that precede artificially long exons are not spliced well. In studies aimed at understanding the importance of exon size on splicing, Berget and colleagues found that precursor RNAs with elongated second exons were deficient for splicing and complex formation. Similar substrates that contained a 5' splice site at the 3' end of the second exon spliced normally, as long as the distance between the 3' and 5' splice sites of the second exon was less than 300 nucleotides (Robberson et al., 1990). In other experiments, mutations in a 5' splice site which decrease complementarity to U1 snRNA resulted in skipping of the preceding exon: intron definition would only predict skipping of that intron containing the mutated 5' splice site (Talerico and Berget, 1990).
**Figure 4.** Mechanisms of substrate recognition. 

**A. Intron Definition.** In this model of spliceosome assembly, interactions between splicing factors span the intron. The SR protein SC35 associates via protein-protein interactions with the splicing factors U2AF and U1 70K, which also contain regions rich in serine and arginine residues. This communication between factors bound at the 5' splice site and those bound at a downstream 3' splice site define the intron as the unit of recognition. 5' ss, 3' ss, 5' and 3' splice sites respectively; BS, branch site; (Py)n, polypyrimididine tract. 

**B. Exon Definition.** The interaction of splicing factors across exon sequences are described by this model. U1 snRNP bound to a downstream 5’ splice site is proposed to facilitate the association of U2AF with the upstream 3’ splice site (see text). Thus, in this model, the formation of the A complex is affected by U1 snRNP binding to the downstream intron. 

**C. Exon Enhancers.** Sequences that stimulate adjacent splice site activity have been defined as exon enhancers. These cis-acting elements have been demonstrated in some cases to bind splicing factors including SR proteins and U1 snRNP. These sequences may substitute for a downstream 5’ splice site in facilitating complex assembly.
In the exon definition model, interactions between splicing factors span the exon: factors bound at the 3' splice site interact with those at the downstream 5' splice site (Figure 4B; Robberson et al., 1990). The splicing machinery searches for a pair of closely spaced splice sites within an exon. The exon is then defined by the binding of both U1 and U2 snRNPs along with splicing factors including U2AF, SC35, and other SR proteins. In the subsequent step, neighboring exons must be juxtaposed to facilitate the actual splicing reactions (for review, see Berget, 1995). This model applies to the splicing of multiple, short exon substrates, common in vertebrates. In yeast, where many pre-mRNAs contain only one intron, and in the first and last exons of larger metazoan pre-mRNAs, other mechanisms need to be invoked.

The exon definition model is appealing in that shorter exons would more easily allow these interactions, while longer exons might prevent the factors from finding one another. The most compelling evidence to the existence of exon definition is that mutations in a 5' splice site often do not only affect the splicing of that intron, but also affect the splicing of the upstream intron. In support of this model, recent studies have demonstrated that U1 snRNPs bound at a downstream 5' splice site could recruit the essential factor U2AF to a weak upstream polypyrimidine tract (Hoffman and Grabowski, 1992). U1 snRNPs primary role may therefore be to promote exon definition by binding to a downstream 5' splice site and facilitating the assembly of the complex A on the upstream 3' splice site.

Another level of splicing regulation is dictated by exon enhancers: sequences that stimulate adjacent splice site activity. These sequence elements may be required to facilitate the splicing of large exons and those which contain sub-optimal splice sites, and they may substitute for a downstream 5' splice site in the exon definition model.

Exon Enhancers
A common characteristic of regulated exons is the presence of sub-optimal splice sites. These weak exon recognition signals probably allow for positive regulation of exon inclusion by slowing down the constitutive splicing process. Exon inclusion in these cases appears to be driven by special sequences that may be recognized by tissue specific or development specific splicing factors (Berget, 1995). Such sequences, referred to as splicing enhancers, have been identified in several regulated exons (Figure 4C; Lavigueur et al., 1993; Sun et al., 1993; Tian and Maniatis, 1993; Watakabe et al., 1993; Xu et al., 1993; Dirksen et al., 1994; Tanaka et al., 1994; Ramchatesingh et al., 1995). These elements are analogous in some ways to the more traditional enhancers that promote transcription and many share common characteristics: many are orientation independent; some act at a distance of up to several hundred nucleotides; all are proposed to bind trans-acting factors. Several examples of the better characterized exon enhancers and the proposed mechanism of action are summarized below.

**Bovine FP**

The terminal exon of the bovine growth hormone pre-mRNA contains a stretch of purine rich sequences that are required for efficient excision of the preceding intron (intron D; Sun et al., 1993a, 1993b). This sequence, designated the FP element, functioned in either orientation to promote the inclusion of terminal exon. UV cross-linking studies showed that the ASF/SF2 bound to the FP element specifically, while SC35 did not. It was further observed that ASF/SF2 stimulates in vitro splicing of this intron. Interestingly, the splicing effect of ASF on the bovine intron was inhibited by the addition of hnRNP A1, reminiscent of hnRNP A1's antagonism of alternative 5' splice selection by ASF/SF2 (Sun et al., 1993b). These results demonstrate that the alternative splicing of this gene may be regulated by the relative levels of ASF/SF2 and hnRNP A1 in vivo.
Furthermore, the requirement for the exon element in the inclusion of exon 5 was diminished when the upstream 5' splice site of intron D was mutated toward U1-consensus (Dirksen et al., 1994). Therefore, dependence on the element is likely to be the result of a poor 5' splice site in the intron. These observations support the proposal that exon enhancers act by facilitating complex assembly around the upstream intron.

**Fibronectin ED1**

The inclusion of the ED1 exon of fibronectin requires the presence of an 81 nucleotide sequence within the exon. Extensive deletion and replacement studies demonstrated that a stretch of 9 nucleotides within this element, GAAGAAGAC, was sufficient to promote splicing in vitro. This sequence facilitates splicing when placed within the exon, in a position independent manner, but only as far away as 293 nt from the 3' splice site. Studies to understand the mechanism of action of this enhancer showed that its presence markedly stimulated the use of the upstream 3' splice site probably by allowing a more efficient interaction of the U2 snRNP with branch site sequences. RNAs containing the repeat specifically inhibited the activity of the element, suggesting that trans-acting factors mediate the activity. These trans-acting factors may include members of the SR family, as SR proteins prepared from calf thymus interacted specifically with the element by gel shift analysis (Lavigueur et al., 1993). Whether the SR proteins bound to the element act by facilitating the association of U2AF to the branch site, in a manner similar to U1 snRNP, is unclear.

**Cardiac Troponin T**

In studying the splicing of the preprotachykinin (PPTK) pre-mRNA, Hoffman and Grabowski showed that the binding of U1 snRNP to the downstream 5' splice site facilitated the association of U2AF with the 3' splice site (Hoffman and Grabowski, 1992). More recently, Grabowski and colleagues asked what sequence substitutions in
the middle exon of PPTK could promote inclusion to an exon containing a suboptimal 5' splice site. Selected sequences could then be classified as exon enhancers, substituting for the U1 snRNP-mediated process. Exon replacements from the troponin T gene (exon 16) and the adenovirus major late transcription unit promoted middle exon selection. Focusing on the troponin T gene, they located a sequence in exon 16 that acted as an exon enhancer. The binding of U2AF to the polypyrimidine tract upstream of the middle exon was enhanced in substrates containing this troponin T element. The stimulation in U2AF binding was dependent on additional nuclear factors and correlated with an increase in exon inclusion. Immunoprecipitation experiments further revealed that SC35 was associated with the enhancer. These data are consistent with the model that factors bound at an exon enhancer facilitate the association of U2AF with the 3' splice site. U2AF is therefore an indirect target of the exon enhancer (Wang et al., 1995; Figure 4C).

The cardiac troponin T gene contains a second enhancer, this one located within exon 5 (Ramchatesingh et al., 1995). The enhancer activity was sensitive to changes in a nine nucleotide purine-rich sequence, GAGGAAGAA. Mutations in this sequence, even conservative purine substitutions, interfered with enhancer function. Again, trans-acting factors have been implicated in enhancer activity. Four SR factors, ASF/SF2, SRp40, SRp55, and SRp75 bound to enhancer RNA by UV cross-linking, while two other, SC35 and SRp20 were not detected in association with this sequence. The altered binding of these SR proteins to mutated element correlated with the reduced activity of the enhancer. While it is not clear whether SR proteins directly affect the activity of this enhancer, the data are consistent with the enhancer model.

**Mouse immunoglobulin μ (IgM)**

A sequence found in the 5' portion of exon M2 in the IgM gene was determined to be important for inclusion of that exon (Watakabe et al., 1993). The element was purine rich, but unlike other enhancers, it did not function in the reverse orientation.
Interestingly, U1 snRNA was associated with this sequence, as evidenced by its specific cross-linking to the element. Further, a consensus 5' splice site could efficiently substitute for the native purine element, while a mutated 5' splice site could not. Thus, some exon enhancers might act by mimicking 5' splice sites to facilitate exon definition.

5' SPlice Site Selection

It is not clear how splice sites are accurately selected among the large pool of potential sites that lie within a pre-mRNA. Modulation of U1 snRNP binding to the pre-mRNA in many cases dictates the 5' splice site choice (for review, see Valcarcel et al., 1995). Specific splicing factors, including members of SR family, influence U1 snRNP binding to 5' splice sites to either facilitate or repress spliceosome assembly, while U6 snRNA proofreads the cleavage site prior to the first reaction. The 5' splice site is therefore recognized twice, by U1 and U6 snRNAs, to ensure that the correct site is used.

When two competing 5' splice sites are present in the pre-mRNA, the spliceosome must decide where to perform the first catalytic step. Generally, the site closer to consensus, with a greater base-pairing potential to U1 snRNA, is chosen as the splice site. U1 snRNP binding, however, is not sufficient to define the site of cleavage for those pre-mRNAs with two similar sites. U1 snRNP was observed to bind indiscriminately to these sequences, with both sites occupied simultaneously on most pre-mRNAs (Eperon et al., 1993). When both sites were occupied, splicing was observed almost exclusively at the downstream 5' splice site. Splicing to an upstream site would thus require that the second site be weaker (have a lower affinity for U1 snRNP), or that the binding of U1 snRNA be weakened at both sites.

The proximity effect of ASF/SF2 on splice site selection was also explained by this double occupancy model (Eperon et al., 1993). ASF/SF2 appeared to improve the binding of U1 snRNP to all competing 5' splice sites, thus promoting the utilization of the
downstream site. ASF/SF2 may also affect 5' splice site selection by directing U1 snRNP to specific sites (Zuo and Manley, 1993; Kohtz et al., 1994). In other studies, the SR proteins SRp40 and SRp55 promoted distal site usage in some pre-mRNAs. These effects were explained by the selective recruitment of U1 snRNP to the distal 5' splice site; this was observed by UV cross-linking (Zahler and Roth, 1995). SR proteins could therefore affect splice site choice by either promoting the binding of U1 snRNP to a specific site, or promoting U1 snRNP occupation of all possible splice sites with the proximal being selected by some other mechanism.

In the absence of U1 snRNP, the 5' splice site is probably recognized by the U6 snRNA. Complex assembly can occur without U1 snRNP, and spliceosome formation most likely relies on the binding of SR proteins which in turn facilitates U2 snRNP and possibly U4/U5/U6 tri-snRNP binding. Accurate 5' splice site selection was maintained in U1 snRNP-depleted extracts, indicating that there exist additional mechanisms for splice site selection.
REFERENCES


CHAPTER II

COMPLEMENTATION BY SR PROTEINS OF PRE-mRNA SPlicing
REACTIONS DEPLETED OF U1 snRNP

This chapter originally appeared in Science 265:1866-1869 (1994).
Individual snRNPs U1, U2, and U4/U6 have been removed from nuclear extracts of HeLa cells by antisense affinity depletion. Addition of a highly purified preparation of SR proteins fully restored splicing activity in reactions depleted of U1 snRNP, but did not reconstitute splicing in reactions depleted of the other snRNPs. Affinity selection experiments revealed that spliceosomes lacking U1 snRNA formed in the U1 snRNP-depleted reactions reconstituted with SR proteins. Thus, high concentrations of SR proteins facilitate the assembly of pre-mRNA into a spliceosome in the absence of interactions with U1 snRNP.
Pre-mRNA splicing takes place within a large complex termed the spliceosome, which contains four small nuclear ribonucleoprotein particles (U1, U2, U4/U6, and U5 snRNPs) and also many non-snRNP protein factors (1). SR proteins belong to a family of non-snRNP splicing factors that are highly conserved from *Drosophila* to primates and contain extensive repeats of the diaminoo acid sequence serine-arginine (2). Proteins of the SR family have been implicated both in constitutive splicing and in the regulation of alternative splicing (3). Recent work suggests that SR proteins function during an early step in commitment of a substrate to splicing (4), facilitate the binding of U1 snRNP to the 5' splice site (5), and also bridge interactions between the 5' splice site and branch site (6). Here we show that high concentrations of SR proteins circumvent the requirement for U1 snRNP in the pre-mRNA splicing reaction.

Antisense 2'-O-methyl oligoribonucleotides were used to deplete HeLa cell nuclear extracts of either U1, U2, or U4/U6 snRNPs (7). These extracts were not active for splicing of added pre-mRNA, but complemented each other in any pair-wise combination (Fig. 1D) (7, 8). Depleted nuclear extracts were assayed for the presence of snRNAs by Northern hybridization (Fig. 1A). In the reaction depleted of U1 snRNP, the amount of this snRNA was reduced approximately to a thousandth of the previous amount, whereas the levels of the non-targeted snRNAs were not greatly affected.

Fractions containing SR proteins complemented splicing of a β-globin substrate in reactions depleted of U1 snRNP (Fig. 1D). Preparations of purified SR proteins were derived from nuclear extracts of HeLa cells as described (2) (Fig. 1B). A typical SR preparation also contained a low level of contaminating snRNAs that were degraded by digestion with micrococcal nuclease. The resulting preparations contained no detectable snRNAs when analyzed by Northern hybridization (Fig. 1C).

After digestion with micrococcal nuclease, SR preparations retained the ability to restore splicing to reactions depleted of U1 snRNP (Fig. 1D). This resistance contrasted with the sensitivity of both nuclear extract and U2 snRNP-depleted extract, which did not
Figure 1. SR proteins restore splicing to U1-snRNP depleted reactions. (A). Antisense affinity depletion of snRNPs from HeLa cell nuclear extracts. RNA recovered from snRNP-depleted nuclear extracts was analyzed by Northern hybridization with snRNA-specific riboprobes. Lane 1 shows a mock-depleted nuclear extract. Lanes 2, 3 and 4 show extracts depleted of U1, U2, and U4/U6 snRNPs, respectively. (B). SR proteins were isolated from HeLa cells, separated by SDS-PAGE, and detected with Coomassie Blue. (C). The SR protein preparation was treated with micrococcal nuclease to degrade endogenous snRNAs (20). RNA recovered from the digested SR preparation was analyzed with snRNA-specific riboprobes as in (A) (lane 4). Lane 2 shows a micrococcal nuclease-digested U2 snRNP depleted nuclear extract. Lanes 1 and 3 show mock-nuclease treated U2 snRNP depleted extract and SR protein preparation, respectively. (D) Splicing of \( \beta \)-globin pre-mRNA is reconstituted in U1-snRNP depleted reactions supplemented with micrococcal nuclease-treated SR proteins (21). Reactions containing U1 snRNP-depleted extract were incubated with mock-nuclease treated SR proteins (lane 11) or nuclease-treated SR proteins (lane 12). Lanes 2, 4, and 7 show reactions with the mock, U1 and U2 snRNP-depleted nuclear extracts, respectively. Reactions containing different combinations of nuclear extracts and/or SR proteins are as indicated. SR, HeLa cell SR proteins; MN, pre-treated with micrococcal nuclease. Splicing intermediates and products are indicated by icons on the left of the panel. The 5' exon species is not shown although it was detected in all active splicing lanes. Lane M shows end-labeled Msp1 fragments of pBR322. Lane 1 shows a reaction containing a mock-depleted nuclear extract that was incubated on ice.
restore splicing to a U1 snRNP-depleted reaction upon nuclease treatment (Fig. 1D). Nuclear extract and U2 snRNP-depleted extract probably complement the absence of U1 snRNP by contributing functional U1 snRNPs; degradation of the U1 snRNA would be expected to render these preparations inactive. This suggests that the activity of the SR preparation is not the result of residual U1 snRNA. The effects of SR proteins were specific, because their addition had little effect on the extent of splicing of mock-depleted reactions and, in addition, did not restore splicing to either U2 or U4/U6 snRNP-depleted reactions (8). In titration experiments, the concentration of SR proteins required to complement the U1 snRNP depleted reactions was approximately tenfold higher than that of the endogenous concentration (8).

The activity of SR preparations complementing a U1 snRNP-depleted reaction was dependent on the particular pre-mRNA. The β-globin substrate was the most active under these conditions whereas an adenovirus pre-mRNA substrate, Ad1, was also spliced (Fig. 2A), but at approximately 10% of the level observed in the mock-depleted reaction. In the case of the Ad1 substrate, addition of SR proteins to the mock-depleted reaction resulted in the activation of a cryptic 5' splice site that is located 125 nucleotides downstream of the normal splice site (Fig. 2B). Both sites were also active in the U1 snRNP-depleted reaction supplemented with SR proteins. A third substrate, pPIP85A, which is largely based on sequences in the Ad1 pre-mRNA but with many sequence variations, was not active for splicing upon addition of SR proteins to U1 snRNP-depleted reactions (8). Thus, there is a sequence specificity for splicing in the absence of U1 snRNP and in the presence of high concentrations of SR proteins.

To investigate further the role of SR proteins in the facilitation of splicing in U1 snRNP-depleted reactions, spliceosomes were formed in reactions containing 0.2 μg substrate RNA that had been synthesized with biotinylated UTP (9). This amount of substrate corresponded to at least a 100-fold excess over the minute levels of U1 snRNA in the depleted reactions. Biotinylated β-globin pre-mRNAs were incubated under
Figure 2. SR proteins restore splicing to an adenovirus pre-mRNA in U1 snRNP-depleted reactions and also activate a cryptic 5' splice site. (A). Splicing reactions containing the Ad1 pre-mRNA substrate supplemented with SR proteins pretreated with micrococcal-nuclease (21). Lanes 2 and 3 show reactions in a mock-depleted nuclear extract with (lane 3) or without (lane 2) added nuclease-treated SR proteins. Lanes 4 and 5 show the corresponding reactions in a U1 snRNP depleted nuclear extract. Icons indicate the migration of splicing intermediates and products generated by selection of either the wild-type or cryptic splice sites. Species marked with a star are those of the cryptic splicing pathway. (B) Comparison of the wild-type and cryptic 5' splice sites in Ad1 pre-mRNA (22). The cryptic 5' splice site region shares sequence similarity with a consensus sequence, although it has a few differences.
A

SR:  

Mock  ΔU1

B

Cryptic 5'SS  CAUCUG / GUCAGAAAAAGACAAU
WT  5'SS  GUUGGG / GUGAGUACUCCCUCU
splicing conditions for 40 minutes in mock, U1 snRNP-depleted, and U2 snRNP-depleted reactions either with or without added SR proteins; the extent of splicing was assayed by addition of a trace amount of labeled substrate (Fig. 3A).

The spliceosomes from these reactions were recovered by chromatography on streptavidin agarose beads, and the bound RNA was analyzed by Northern blot analysis (Fig. 3B). All five snRNAs were associated with the spliceosome in reactions containing mock-depleted extract. The addition of excess SR proteins had a small stimulatory effect on splicing, and also on the amount of snRNAs in spliceosomes. Trace levels of U2, U4, U5 and U6 snRNAs were affinity selected in the U1 snRNP-depleted reactions (Fig. 3B), consistent with the very low level of splicing that was observed under these conditions. Addition of excess SR proteins, which restored splicing to mock depleted reaction levels (Fig. 3A), enhanced the binding of U2 snRNA to levels found in the mock reactions. Furthermore, U4, U5, and U6 snRNAs were also enriched in complexes formed under these conditions. No U1 snRNA above background was selected in the U1 snRNP-depleted reaction upon addition of high concentrations of SR proteins (Fig. 3B).

Analysis of the U2 snRNP-depleted reactions revealed that a small amount of U1 snRNA was associated with the substrate in the absence of additional SR proteins, and that this amount increased upon addition of SR proteins (Fig. 3B). However, consistent with the lack of splicing under these conditions (Fig. 3A), the addition of SR proteins did not promote the association of the other snRNAs (Fig. 3B). Similar results to these and the above were obtained from selection experiments carried out using the Ad1 pre-mRNA substrate (10).

Pre-mRNA splicing is efficient in reactions depleted of U1 snRNP if the concentration of SR proteins is increased 10-fold. We suggest that splicing under these conditions is independent of the activity of U1 snRNP. This conclusion is also supported by the absence of detectable U1 snRNA in affinity-selected spliceosomes under conditions in which other spliceosomal snRNAs were clearly present. Furthermore,
Figure 3. Spliceosomes lacking U1 snRNA form on the β-globin pre-mRNA during splicing. (A) Splicing reactions were monitored by following trace amounts of radiolabeled β-globin substrate, co-incubated with a near-saturation level of biotinylated-cold pre-mRNA. Reactions containing mock depleted nuclear extract (lanes 1 and 2), U1 snRNP-depleted nuclear extract (lanes 3 and 4), or U2 snRNP-depleted nuclear extract (lanes 5 and 6) were incubated with (lanes 2, 4 and 6) or without (lanes 1, 3 and 5) SR proteins pre-treated with micrococcal nuclease. The 5' exon species is not shown although it was detected in all active splicing lanes. (B) Spliceosomes formed in the reactions shown in (A) were affinity selected on streptavidin agarose beads (9). Co-selected snRNAs were detected by Northern hybridization as in Fig.1A. Lane 1 contains RNA extracted from 10μl nuclear extract as a marker. Lane 2 shows RNA affinity selected from a reaction containing no substrate RNA.
A

SR: - - + - + - +

Mock  ∆U1  ∆U2

[Diagram with lanes labeled 1 to 6]
Fraction: SR:

<table>
<thead>
<tr>
<th></th>
<th>Mock</th>
<th>ΔU1</th>
<th>ΔU2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

1 2 3 4 5 6 7 8

U2 U1 U4 U5 U6
substrate RNAs that are processed with similar efficiencies in U1 snRNP-dependent reactions have differential activities in reactions depleted of U1 snRNP and supplemented with SR proteins. Even with these strong indications, it is impossible to prove chemically that the minute level of U1 snRNP present in the depleted reaction is not active in a very transient mode in the promotion of splicing. However, we propose that these results shift the argument to the premise that U1 snRNP is not required for splicing under all conditions.

That high concentrations of SR proteins permit processing of pre-mRNA in reactions depleted of U1 snRNP suggests that U1 snRNA is not essential for catalysis of either of the transesterification steps in splicing. This finding was anticipated by previous studies which show that U1 snRNP was easily displaced from the spliceosome (11), and the lack of an obvious analog to U1 snRNA in trans-splicing reactions (12). Further evidence that a spliceosome can form without U1 snRNP is suggested by recent studies showing that spliceosome-type complexes containing U2, U5 and U4/U6 snRNAs assemble on short oligoribonucleotides encompassing a 5' splice site (13). Formation of these complexes was dependent on the recognition of the 5' splice site sequence, but did not depend on the presence of an intact U1 snRNA. Consistent with the above, it was recently reported that SR proteins reconstituted splicing reactions in which U1 snRNP was debilitated by site-specific blockage of the 5' end of U1 snRNA. In these reactions, no U1 snRNA was detected interacting with the 5' splice site by psoralen cross-linking (14).

The 5' splice site sequence is recognized by components of the spliceosome other than U1 snRNA. Recent experiments have revealed a U6 snRNA-5' splice site interaction by detection of allele-specific complementation between mutations in the 5' splice site and mutations in U6 snRNA, and also by ultraviolet light-induced cross-links between these two RNAs (15). Consistent with specific recognition of the 5' splice site sequence by spliceosome factors other than U1, we have found that mutations at either
+1, +2, +5, or +6 in the 5' splice site region have similar effects on reactions containing either a mock-depleted extract, or U1 snRNP depleted extract reconstituted with SR proteins (16). Since bases at these positions are certainly recognized in a rate limiting step by U1 snRNA under normal conditions, the same bases must also be recognized in another potentially rate limiting step post-interaction by U1 snRNP, probably by U6 snRNA.

The mechanism by which SR proteins bypass the requirement for U1 snRNP is unclear. U1 snRNP recognizes the 5' splice site by sequence complementarity (17), and promotes commitment to the splicing pathway by stabilizing the binding of U2 snRNP to the branch region (7, 18). The SR subdomain of the U1 snRNP 70 kD protein may interact either directly or indirectly with the SR domain of the U2AF protein which is required for the stable binding of U2 snRNP (6). We propose that this U1 snRNP-mediated signal for U2 snRNP binding and spliceosome formation is not essential under conditions of high SR proteins, because the SR proteins bind the substrate directly and stabilize the subsequent association of U2 snRNP. This suggestion is consistent with the observed differences in splicing activity of specific substrates in the U1 snRNP-depleted-high SR protein reactions. SR proteins are thought to bind RNA in a sequence specific fashion (19), and thus substrate pre-mRNAs might be expected to have differential activity in reactions dependent upon recognition by SR proteins. Further studies on the sequence specificities of reactions in the presence of high levels of SR proteins will probably be informative in determining the mechanisms by which SR proteins function to promote splicing in the absence of U1 snRNP.
REFERENCES AND NOTES


7. S. M. L. Barabino, B. J. Blencowe, U. Ryder, B. S. Sproat, A. I. Lamond, Cell 63, 293 (1990). Nuclear extracts were depleted of specific snRNPs as described in: [B. J. Blencowe, S. M. L. Barabino, in Methods in Molecular Biology (Humana Press, USA, 1994), vol. 37, In Press]. Mock depleted extracts were carried through the depletion protocol in the absence of an antisense oligonucleotide. RNA from nuclear extracts was recovered by proteinase K treatment, followed by phenol extraction and ethanol precipitation, prior to electrophoresis in urea-polyacrylamide gels. The biotinylated
antisense oligonucleotides were synthesized on an Applied Biosystems Oligonucleotide Synthesizer with commercially available phosphoramidites.

8. J. D. Crispino and P. A. Sharp, unpublished data.

9. Biotinylated RNA was transcribed as described [P. J. Grabowski and P. A. Sharp, Science 233, 1294 (1986)]. 100 µl splicing reactions containing 0.2 µg β-globin substrate were incubated for 40 min. at 30°C. Immediately prior to incubation at 30°C, a small amount of labeled pre-mRNA was added to 10 µl of each splicing reaction. The remaining 90 µl of each reaction was incubated at 30°C and then mixed for 1 hour at 4°C with pre-blocked streptavidin agarose beads (Sigma). The beads were washed several times using a 350 mM KCl buffer as described [U. Ryder, B. S. Sproat, A. I. Lamond, Nucleic Acids Research 18, 7373 (1990)]. Affinity selected RNA was released by proteinase K treatment, phenol/chloroform extracted, then recovered by ethanol precipitation for analysis by Northern hybridization.

10. J. D. Crispino and P. A. Sharp, unpublished data.


16. J. D. Crispino and P. A. Sharp, unpublished data.


20. SR proteins were prepared from HeLa cells as described (2) but with a second MgCl$_2$ precipitation. To remove endogenous snRNAs, SR proteins and nuclear extract were treated with micrococcal nuclease for 30 minutes at 30°C, in the presence of 1 mM CaCl$_2$. The extent of degradation of the snRNAs was assayed by Northern hybridization with snRNA-specific riboprobes. The amounts of U2 snRNP-depleted extract and SR proteins digested corresponded to the amounts required to obtain equivalent levels of complementing activity, as assayed by the ability to restore splicing to a U1 snRNP-depleted extract. After the incubation with nuclease, EGTA was added to a final concentration of 3 mM.

21. Splicing reactions were performed essentially as described [P. J. Grabowski, R. A. Padgett, P. A. Sharp, *Cell* **37**, 415 (1984)]. Splicing-complementation assays contained
30% nuclear extract, supplemented with 20% of a second extract or 1 µg SR proteins. In complementation controls lacking extract or SR proteins, buffer D [J. D. Dignam, R. M. Lebowitz, R. G. Roeder, *Nucl. Acids. Res.* 11, 1475 (1983)] containing 5% glycerol was added instead. Splicing reactions were incubated for two hours.

22. RNA species from splicing reactions with the Ad1 pre-mRNA were extensively characterized by comparison of the mobility of the RNA products, and by cDNA cloning of the exon-product RNAs after elution from polyacrylamide gels. The cDNA and second-strand synthesis was performed with rTh polymerase (Perkin-Elmer Cetus) in coupled reverse transcription-PCR. The cDNAs were inserted into pBS-, and the DNAs of several clones were sequenced.

23. The authors thank R. Cook and S. Schultz of the MIT Biopolymers Facility for synthesizing the biotinylated 2'-O-methyl oligoribonucleotides. R. Issner and Y. Qiu are thanked for their excellent technical assistance. A. MacMillan, P. McCaw, J. Pomerantz and C. Query kindly made helpful comments on the manuscript. Many thanks also to M. Siafaca for secretarial support. The work was supported by United States Public Health Service grant RO1-GM34277 and RO1-AI32486 from the National Institute of Health to P.A.S. and partially by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute. B.J.B. was supported by a Human Frontiers Science Program Organization Long-Term Fellowship.
CHAPTER III

A U6 snRNA:PRE-mRNA INTERACTION CAN BE RATE-LIMITING FOR U1-INDEPENDENT SPLICING.

ABSTRACT

The full set of consensus sequences at the 5' splice site is recognized during splicing of pre-mRNA in extracts depleted of U1 snRNP. High concentrations of HeLa SR proteins, or purified SC35 alone, promote the splicing of specific RNA substrates, bypassing the requirement for U1 snRNP in formation of the U2 snRNP–pre-mRNA complex. Under these conditions, mutations in the substrate which increase the sequence complementarity between U6 snRNA and the 5' splice site region can facilitate splicing. This provides additional strong evidence that U1 snRNP is not essential for splicing. Thus, the consensus sequence at the 5' splice site is probably recognized twice during splicing of most introns; however, some pre-mRNAs could potentially be processed in the absence of interactions with U1 snRNP in regions of the nucleus containing high concentrations of SR proteins.
INTRODUCTION

Pre-mRNA splicing occurs within a large multi-subunit complex termed the spliceosome. The spliceosome, which is composed of four snRNP particles (U1, U2, U4/6 and U5) and many non-snRNP protein factors, forms in a step-wise manner on the substrate (for review, see Moore et al., 1993; Madhani and Guthrie, 1994). In the first commitment step, U1 snRNP, in conjunction with the splicing factor U2AF and several members of the SR protein family, associates with the pre-mRNA to form the E complex (Michaud and Reed, 1993; Staknis and Reed, 1994). During this reaction, U1 snRNP binds to the 5' splice site by base pairing (Zhuang and Weiner, 1986; Séraphin et al., 1988; Siliciano and Guthrie, 1988) and promotes the binding of U2 snRNP to the branch region, forming the A complex (Konarska and Sharp, 1986, 1987; Séraphin and Rosbash, 1989; Barabino et al., 1990; Michaud and Reed, 1991). Finally, the complete spliceosome is formed by the addition of the U4/5/6 tri-snRNP.

Several rearrangements occur within the spliceosome before the first covalent reaction resulting in the cleavage of the 5' splice site and branch formation. First, U4 snRNA is separated from the U6 snRNA so that U2 and U6 can interact; these two RNAs are proposed to form the catalytic center (Madhani and Guthrie, 1992; McPheeters and Abelson, 1992). In addition, the U1 snRNA is probably displaced from the 5' splice site (Konforti et al., 1993), while U6 snRNA binds to distal nucleotides in the 5' splice site (+4 through +6). The two RNAs can be cross-linked in this region by UV irradiation (Sawa and Abelson, 1992; Sontheimer and Steitz, 1993). Further, allele specific complementation between mutations in the 5' splice site and in U6 snRNA have demonstrated a base-pair interaction between these RNAs at position +5 of the intron and further suggest that U6 snRNA can pair with additional nucleotides downstream of the +5 position (Kandels-Lewis and Séraphin, 1993; Lesser and Guthrie, 1993; Sun and Manley, 1995).
The distal sequences in the 5' splice site are therefore recognized at least twice during the splicing reaction, by the U1 and U6 snRNAs. The other nucleotides in the 5' splice site (-1 through +3) may also be recognized a second time, probably in the binding of the U4/5/6 tri-snRNP. Formation of spliceosome-type complexes containing U2, U5, and U4/U6 snRNAs on short oligoribonucleotides comprising a 5' splice site occurs when sequences at the 5' terminus of U1 snRNA are blocked (Konforti et al., 1993; Konforti and Konarska, 1994). In these experiments, nucleotides at positions -2 through +5 were critical for both efficient binding of the tri-snRNP and for UV cross-linking of the U6 snRNA to the RNA oligo.

SR proteins, which contain extensive repeats of the dipeptide sequence serine-arginine, have been implicated in constitutive splicing and in the regulation of alternative splicing, such as the sex-determination pathway in Drosophila (Ge et al., 1991; Krainer et al., 1991; Fu et al., 1992; Zahler et al., 1992; Tian and Maniatis, 1993; Lynch and Maniatis, 1995). These proteins function during an early step in the commitment of a substrate to splicing (Fu, 1993), in the formation of the E complex (Staknis and Reed, 1994), and in bridging interactions between the 5' splice site and the branch site (Wu and Maniatis, 1993). Addition of SR proteins can switch the utilization of two competing 5' splice sites (Zahler et al., 1993; Zahler and Roth, 1995), possibly by directly binding to 5' splice sites and recruiting U1 snRNP (Kohtz et al., 1994; Zuo and Manley, 1994; Zahler and Roth, 1995). SR proteins also participate in splicing by binding to exon splicing enhancers: sequences that stimulate adjacent splice site activity (Lavigueur et al., 1993; Sun et al., 1993; Tian and Maniatis, 1993; Watakabe et al., 1993; Xu et al., 1993; Dirksen et al., 1994; Tanaka et al., 1994). SR proteins bound at these enhancers are postulated to promote U2 snRNP binding to the branch site, probably by recruiting and stabilizing the interaction of the branch site region with U2AF (Wang et al., 1995).

We previously demonstrated that SR proteins promote efficient splicing in extracts that have been depleted for U1 snRNP (Crispino et al., 1994). In a separate
study, SR proteins were shown to reconstitute the splicing of pre-mRNAs in extracts where the U1 snRNP was blocked by an antisense oligonucleotide (Tarn and Steitz, 1994). SR proteins have also been demonstrated to promote the binding of U2 snRNP to the branch site of an RNA lacking a 5' splice site (Tarn and Steitz, 1995).

Here, we present evidence that splicing in U1 snRNP-depleted extracts supplemented with SR proteins has a rate-limiting step distinct from that in mock-depleted extracts. In U1-independent reactions, a rate-limiting step is the base pairing of the pre-mRNA with U6 snRNA, probably during the binding of the U4/5/6 tri-snRNP. We also demonstrate that the entire 5' splice site consensus sequence is recognized in the U1 snRNP-depleted reactions.
RESULTS

SR proteins promote complex formation in U1 snRNP-depleted extracts in a substrate specific reaction

The efficiency with which SR proteins complemented a U1 snRNP-depleted reaction was dependent on the particular pre-mRNA (Crispino et al., 1994). The β-globin pre-mRNA was processed in the U1 snRNP-depleted extract supplemented with SR proteins (AU1+SR) to the same extent as in mock-depleted extracts, while an adenovirus pre-mRNA substrate, Ad1, was processed at approximately 10% of the level observed in the mock-depleted extract. PIP85.A, a pre-mRNA with sequences based largely upon Ad1, was not active for splicing in AU1+SR reactions. To further analyze the nature of this substrate specificity, splicing reactions containing either the β-globin or PIP85.A pre-mRNAs were assayed for both splicing and spliceosome formation. As expected, the AU1+SR reactions spliced β-globin to the same extent as the mock-depleted extract (Figure 1A, compare lanes 1 and 2 with 7 and 8) while the PIP85.A substrate was inactive (Figure 1C, lanes 16-20).

Analysis of the reactions in figures 1A and 1C by native gel electrophoresis (Konarska and Sharp, 1986, 1987) revealed the predicted complexes in the mock-depleted extracts (Figures 1B and 1D, mock lanes). Complex A, composed of U2 snRNP stably associated with the branch site, and complex B, formed by the binding of the U4/5/6 tri-snRNP to complex A, were rapidly generated. The slowest migrating complex, C, which is produced by rearrangements within the spliceosome and addition of other protein factors, was difficult to distinguish in these gels. Addition of SR proteins to mock-depleted extracts had little effect on complex formation or splicing (Figures 1A through 1D, mock+SR lanes).
Figure 1. SR proteins promote complex A formation in reactions depleted of U1 snRNP. (A) Time course of splicing for β-globin pre-mRNA in mock and U1 snRNP-depleted extracts, supplemented with either 450 ng of HeLa SR proteins or an equal volume of buffer. Splicing intermediates and products are identified by icons on the left. The 5' exon RNA is not shown, although it was detected in all active splicing lanes. The species marked with an asterisk corresponds to a fragment protected by U2 snRNP from exonuclease digestion. (B) Native gel analysis of the complexes formed on the β-globin pre-mRNA during the splicing reactions in figure A. The identity of the H, A, B, and C complexes are indicated to the left. (C) Time course of splicing for PIP85.A pre-mRNA in mock and U1 snRNP-depleted extracts, supplemented with either 450 ng of HeLa SR proteins or an equal volume of buffer. (D) Native gel analysis of complexes formed on the PIP85.A substrate in the reactions of figure C.
U1 snRNP is required for the stable binding of U2 snRNP under normal conditions (Barabino et al., 1990). Consistent with this earlier finding, complex A formation was impaired on either pre-mRNA in the absence of U1 snRNP (Figures 1B and 1D, AU1 lanes). However, addition of high concentrations of SR proteins greatly stimulated the formation of complex A on the β-globin substrate in the U1 snRNP-depleted reactions (Figure 1B, AU1+SR lanes). The slower migrating B and C complexes were also present, consistent with the observed splicing activity in these reactions (Figure 1A, AU1+SR lanes). These complexes are less stable than those formed in mock-depleted extracts, as they were not detectable when reactions were electrophoresed at room temperature instead of at 4°C (data not shown). Addition of SR proteins slightly increased the amount of U2 snRNP bound to the PIP85.A substrate in reactions depleted of U1 snRNP (Figure 1D, AU1+SR lanes); this effect was minimal when compared to that observed with the β-globin substrate.

Thus, the difference between the splicing activities of the two substrates in the U1 snRNP-depleted/high SR conditions is reflected by the efficiency of formation of the U2 snRNP complex. SR proteins stimulated the transition from H complex to complex A in a substrate specific manner. The identities of the complexes formed under the above conditions were confirmed by affinity selection experiments using biotinylated pre-mRNAs (data not shown). In addition, both pre-mRNAs failed to form detectable snRNP-bound complexes in the U2 snRNP-depleted extracts, consistent with the A complex being the earliest detectable species in this assay (data not shown).

**SC35 and SRp55 promote splicing in U1 snRNP-depleted extracts.**

SR proteins derived from HeLa cells contain six major species: SRp75, SRp55, SRp40, ASF/SF2, SC35 and SRp20 (Zahler et al., 1992). To determine whether a single species can reconstitute activity, SC35 and SRp55, produced by baculovirus infection of
Figure 2. Purified SR proteins complement splicing extracts depleted for U1 snRNP. The β-globin pre-mRNA was incubated for 90 minutes in either mock or U1 snRNP-depleted nuclear extracts and supplemented with HeLa SR proteins (900 ng), purified SC35 (100, 300, and 900 ng), or purified SRp55 (100, 300, or 900 ng). The bands and icons are as described in Figure 1.
insect cells, were assayed for their ability to restore splicing to U1 snRNP-depleted extracts (Figure 2). The splicing of the β-globin pre-mRNA in mock-depleted and U1 snRNP-depleted extracts supplemented with different SR proteins was analyzed. Purified SC35 alone was sufficient to restore an equivalent level of splicing to the U1 snRNP-depleted extract as the HeLa SR proteins (compare lanes 4-6 with lane 3). Purified SRp55 also promoted splicing in the depleted extract, albeit with a reduced efficiency (lanes 7-9).

The activity of SC35 also parallels that of the complete SR preparation in its sequence specificity. Although SC35 was efficient in restoring splicing reactions with the β-globin RNA, it did not promote splicing of the PIP85.A substrate in U1 snRNP-depleted extracts and only weakly restored splicing to the Ad1 pre-mRNA (data not shown). SC35 also has the same activity in assays for complex formation as described for the SR proteins above (data not shown). In order to investigate the role of SR proteins in splicing in a more-defined system, purified SC35 was used in all subsequent experiments.

The entire 5' splice site is recognized in U1 snRNP-depleted extracts supplemented with SC35

The sequences of the 5' splice sites are conserved in mammalian introns with a consensus of AG/GUAAGU (the bar denotes the 5' splice site). A subset of these nucleotides have been shown to be important in binding of U1 and U6 snRNAs during the course of the splicing process (for review, see Madhani and Guthrie, 1994). In reactions which lack U1 snRNP, one might predict that those nucleotides whose sole function is in the base pairing with U1 snRNA would be dispensable, while the bases that are involved in additional contacts would be required. In order to investigate the role of the 5' splice site during splicing in the absence of U1 snRNP, point mutations were
Figure 3. The 5' splice site is recognized in the absence of U1 snRNP. (A) Time course of splicing of the wild-type and the +1 G → A point mutant in complete or reconstituted reactions. The complete reactions contained mock-depleted nuclear extract, while the reconstituted reactions were performed with U1 snRNP-depleted extract supplemented with 300 ng of SC35. The bands and icons are as described in Figure 1. For the +1 G → A point mutant, cryptic splice sites are utilized, leading to alternate forms of lariat intermediates and products. (B) Graphical representation of the relatives rates of the first step of splicing. Values are the average of 3 or more experiments with the error bars indicating one standard deviation of the mean, except for the following mutants: -2 G → A, +4 A → U, and +6 U → A where the value represents the average of two experiments, with the error bars the range of the two determinations.
introduced into the β-globin pre-mRNA at several positions. These mutants were then incubated in both mock and ΔU1+SC35 (reconstituted) extracts. The rates of splicing in both reactions were determined and compared with the rates obtained with the wild-type pre-mRNA (see materials and methods).

For the wild-type substrate, the rates of appearance of splicing products in mock and reconstituted reactions were similar (compare Figure 3A mock and ΔU1+SC35 splicing). Parallel splicing reactions with the +1G → A mutant are shown. This mutation has previously been shown to delay 5' exon cleavage and to inhibit splicing by blocking the second step of the reaction (Lamond et al., 1987). In the mock-depleted extract, the RNA proceeded through the first step at approximately 20% the rate of wild-type (Figure 3A). Splicing of the equivalent step for this mutant in the reconstituted reactions was also reduced to 20% of wild-type. This indicates that the G at the +1 position is critical for efficient splicing in both reaction types.

The rates of splicing for several other point mutations are summarized in Figure 3B. Positions -1, +1, +2 and +4 showed significantly reduced splicing rates in both mock and reconstituted reactions. Therefore, these nucleotides are important in both reaction pathways. In contrast, the +5 and +6 mutations had little effect on splicing in mock-depleted reactions, but demonstrated a significant effect on splicing in the reconstituted reactions. Changes in the identity of the -2 position did not affect the rate of the first step in either pathway. This suggests that the nucleotides between the -1 and +2, and +4 through +6 positions are critical in a rate-limiting step of splicing in the U1 snRNP-depleted/SR supplemented reaction.

Second site mutations that increase potential pairing to U6 snRNA stimulate splicing only in the reconstituted reaction.
Several results suggest that the rate limiting step for splicing in the reconstituted reactions is the binding of U6 snRNA to the distal sequences of the 5' splice site. First, mutations in positions +5 and +6, which are predicted to disrupt base pairing with U6 snRNA, have a stronger phenotype on splicing in reconstituted reactions than in normal reactions. Second, two different point mutations at the +4 position had different effects on the rate of splicing in the reconstituted reaction only. The +4 A → U mutant spliced two-fold more efficiently than the +4 A → G mutant. According to the current model of U6 snRNA pairing to the 5' splice site (Figure 4A), the adenosine at position 43 in U6 snRNA could base pair with the +4 A → U, increasing the stability of the helix and, hypothetically, the rate of splicing. This argument is qualified by the fact that the wild-type A at +4 would also not pair with position 43 of U6 snRNA, but is more efficient than the +4U mutation. This may suggest that this position is also recognized by another component of the spliceosome. Interestingly, the same allele specificity was observed in studies of the binding of the U4/5/6 tri-snRNP to a short oligoribonucleotide comprising the 5' splice site (Konforti and Konarska, 1994).

If the binding of U6 snRNA to the 5' splice site is rate-limiting, then mutations that stabilize the base pairing of the 5' splice site with U6 snRNA might accelerate the rate of splicing in the reconstituted reactions. Previous studies have shown that positions +7, +8 and +9 of an intron can be recognized by base-pairing with U6 snRNA (Lesser and Guthrie, 1993; Kandels-Lewis and Séraphin, 1993). Positions +7 and +9 were changed to the complementary bases of the U6 snRNA nucleotides 38 and 40 in the context of the wild-type β-globin 5' splice site and in each of three point mutants (Figure 4A). The rates of splicing of these substrates were analyzed in mock reactions and in SR reconstituted reactions (Figure 4C).

Splicing of the +4 A → G and +5 G → A mutants was dramatically increased when the second site mutations were introduced. An increase from 27% to 100% of wild-type splicing was observed in the reconstituted reactions for the +4 A → G mutant,
Figure 4. U6 snRNA binding is rate-limiting for splicing in the ΔU1+SR reactions. (A) Diagram of the 5' splice site sequence of the β-globin pre-mRNA and the proposed pairing of the human U6 snRNA (for review, see Madhani and Guthrie, 1994). The mutations introduced, which increase the complementarity between the 5' splice site and the U6 snRNA, are shown. (B) Splicing time courses of β-globin, the +4 A → G mutant, the +7 U → A/ +9 G → C mutant, and the double mutant are shown. Purified SC35 was added to the U1 snRNP-depleted extract in the reconstituted reactions. The bands and icons are as described in Figure 1. (C) Graphical representation of the relative splicing rates for the +2, +4, and +5 mutants, along with the respective double mutants. Values are the average of 3 or more experiments, with the error bars indicating one standard deviation of the mean, except for the following mutants: +2 U → C/ +7 U → A/ +9 G → C and the +5 G → A/ +7 U → A/ +9 G → C, where the values represent the average of two determinations, with the error bars the range of the two determinations.
A

GUGAGUUUGGG\[gg\]β-globin

AGACAUA\[42\]GAGAGAUAGCA---U6 snRNA

5' splice site mutants

C

% splicing relative to wild-type

Mock
Reconstituted

5' splice site mutants

7U→A 9G→C
2U→C 2U→C 4A→G 4A→G
7U→A 7U→A 5G→A 5G→A
9G→C 9G→C 7U→A 7U→A
B

Wild-type +4 A→G mutant

<table>
<thead>
<tr>
<th>time (min)</th>
<th>Mock</th>
<th>ΔU1+SC35</th>
<th>Mock</th>
<th>ΔU1+SC35</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 30 40 50 60 75 90</td>
<td>0 30 40 50 60 75 90</td>
<td>0 30 40 50 60 75 90</td>
<td>0 30 40 50 60 75 90</td>
<td></td>
</tr>
</tbody>
</table>

Mock

ΔU1+SC35

+7 U→A
+9 G→C

+4 A→G
+7 U→A
+9 G→C
with little change in the splicing in mock extracts (Figures 4B and 4C). Similarly, the
splicing of the +5 G → A mutant increased from 37% of wild type rate in the
reconstituted reaction to 92% when the base pairing to U6 snRNA was increased. Again,
the double mutants spliced to the same extent in the mock extracts.

In contrast, the +2 U → C mutant was not significantly rescued by the mutations
which increased U6 pairing (Figure 4C). Splicing was improved from 31% to 41%, with
no change in the splicing in mock extracts. Recognition of the +2 U → C mutation is
apparently not limited by base-pairing to U6 snRNA, but by some other interaction.

**U2-pre-mRNA complexes are formed in the reconstituted reactions prior to
recognition of the 5' splice site.**

The above results suggest that association of the U4/5/6 tri-snRNP can be rate-
limiting in the reconstituted reaction. This binding depends upon the presence of the U2
snRNP-pre-mRNA complex, which must have been available in non-limiting amounts
even in reactions with RNAs containing mutations in the 5' splice site. In normal
reactions, the association of the 5' splice site with U1 snRNP is rate-limiting, and a delay
in the rate of formation of the U2 snRNP complex can be detected for substrates with
mutations in the 5' splice site. These two results indicate that mutations in the 5' splice
site should affect formation of the U2 snRNP A complex differently under the two
reaction conditions.

The rate of complex A formation with the +1 G → A mutation was compared to
that of wild-type substrate, both in mock and ΔU1+SC35 reactions (Figure 5). As
previously demonstrated for this mutant (Lamond et al., 1987), formation of the U2
snRNP complex A was greatly reduced compared to wild-type pre-mRNA in mock
reactions. In contrast, for the same mutant, complex A formation in the reconstituted
reactions was decreased only slightly from that for the wild-type substrates. The +4
A → G mutant also showed reduced complex formation specifically in mock-depleted reactions (data not shown). These results indicate that the 5' splice site is recognized post-U2 snRNP binding in U1 snRNP-depleted extracts supplemented with SR proteins and pre-U2 snRNP binding under complete reaction conditions.
Figure 5. Native gel analysis of snRNP complexes formed on wild-type β-globin and the +1 G → A mutant in the course of both complete and reconstituted splicing reactions. Reaction mixtures were treated with heparin prior to loading on the native polyacrylamide gel.
Mock

Wild-type

+1 G→A mutant

<table>
<thead>
<tr>
<th>time (min)</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>5</th>
<th>15</th>
<th>30</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
</tr>
</thead>
</table>

Wild-type

ΔU1+SC35

Mock

ΔU1+SC35
DISCUSSION

High concentrations of SR proteins circumvent the requirement for U1 snRNP in pre-mRNA splicing. This bypass reaction is substrate specific, probably reflecting the efficiency of binding of SR proteins. In this reaction, SR proteins first promote the rapid association of U2 snRNP with the branch site, in a manner independent of either U1 snRNP or the 5' splice site (Figure 6A). The entire 5' splice site consensus sequence is required in a step after the formation of complex A (Figure 6B). This recognition of the 5' splice site occurs in a rate-limiting reaction with, or after, the binding of the U4/5/6 tri-snRNP. The distal nucleotides in the 5' splice site are probably recognized by base pairing with U6 snRNA, while the proximal sequences interact with additional spliceosomal components.

Base pairing with U6 snRNA is rate-limiting in U1 snRNP-depleted extracts

The consensus sequence at the 5' splice site was recognized in the U1 snRNP-depleted/SR reconstituted reactions, since mutations in these sequences reduced the efficiency of splicing. This suggests that most of the positions between -1 and +6 of the intron are recognized by components other than U1 snRNP, probably U6 snRNA and other factors in the spliceosome (Figure 6B). These results are consistent with previous experiments that characterized the binding of a spliceosome-type complex lacking U1 snRNP to oligoribonucleotides containing the 5' splice site sequence (Konforti and Konarska, 1994). They are also consistent with the sequence specificity of trypanosome and nematode trans-splicing, which does not apparently require an analog of U1 snRNP, but does depend upon the presence of the consensus sequence of the 5' splice site (Nilsen, 1993).
Figure 6. (A) Proposed scheme for splicing in complete and SR reconstituted reactions. In complete reactions, U1 snRNP and SR proteins cooperatively interact to form E complex (Staknis and Reed, 1994). U2 snRNP then binds to the branch site to produce complex A. In U1-depleted extracts, excess SR proteins first form a complex with the pre-mRNA, and then promote the association of U2 snRNP with the branch site. (B) Interactions at the 5' splice site during the course of splicing. In complete reactions both U1 and U6 snRNAs base pair to 5' splice site sequences, and an additional factor, X, recognizes the proximal nucleotides in the 5' splice site. In the SR reconstituted reactions, U6 snRNA and factor X bind to the 5' splice site in rate-limiting steps.
A

SRs·RNA → +SRs

RNA → +U2

U2·RNA → +U4/5/6

U2/4/5/6·RNA → Splicing

B

U1

UCAUUCAUGm3

AG

GUAAGU

GUAAGU

GCAUAGCA

AG

GUAAGU

GUAAGU

GCAUAGCA

AG

GUAAGU
Previous evidence suggests that U6 snRNA binds to sequences +4, +5, and +6 of the 5' splice site in the spliceosome (Kandels-Lewis and Séraphin, 1993; Lesser and Guthrie, 1993; Sun and Manley, 1995). Disruption of these interactions could explain the phenotypes of mutations in these positions in the absence of U1 snRNP. The proximal sequences, +1 and +2 of the 5' splice site are also critical for efficient splicing in the absence of U1 snRNA. These sequences must interact with other components of the spliceosome. One of these components is probably the p220 protein of the U4/5/6 particle, which cross-links to the base at the +2 position (J.L. Reyes and M.M. Konarska, personal communication).

Changes in the nucleotides at positions +4 and +5 reduced the efficiency of splicing in both the U1 snRNP-dependent pathway and in the U1 snRNP-depleted/SR reconstituted pathway. The rates of splicing of these mutants were increased significantly in the reconstituted pathway, when additional changes at +7 and +9 which improved the base pairing of U6 snRNA to the substrate were introduced. However, the same changes did not increase the rate of splicing in the U1 snRNP-dependent pathway. This suggests that positions +4 and +5 were recognized by the U6 snRNA in a rate-limiting step of splicing in the former pathway, but not in the latter. In the latter reaction, a step prior to the interaction with U6 snRNA is limited by mutations in the +4 and +5 positions, probably the U1 snRNP-dependent binding of U2 snRNP.

It is interesting that mutations at the +2 position cannot be rescued by increasing sequence complementarity to U6 snRNA. This suggests that this position is recognized in a rate-limiting step other than the interaction with U6 snRNA, perhaps by a factor critical for the chemistry of the first step in splicing.

Although our data do not physically identify U6 snRNA binding as the rate-limiting event in U1 snRNP-depleted reactions supplemented with SR proteins, the combination of the mutant suppression results and the fact that efficient formation of complex A occurs under these conditions strongly implicate the interaction with U6 snRNA.
snRNA as rate-limiting for recognition of the 5’ splice site. The effects of the intron +7 and +9 mutational changes on splicing are most easily explained by the facilitation of the binding of U6 snRNA. Since these positions are not conserved among introns, it is unlikely that some uncharacterized factor specifically binds to these sequences and affects splicing in vitro. The only reasonable model is that U6 snRNA recognizes these positions in a potentially rate-limiting manner in the absence of U1 snRNP.

The data presented here give further strong evidence that the splicing reaction in the extracts depleted of U1 snRNP proceeds in a manner independent of this snRNP. The rate-limiting step in the depleted-reconstituted reaction was recognition of the 5’ splice site by U6 snRNA, while under normal conditions, the rate-limiting step is probably recognition of this sequence by U1 snRNP. Thus, recognition of the 5’ splice site in the U1 snRNP-depleted reactions is probably not mediated by contaminating levels of U1 snRNP, as it is unlikely that reducing the level of U1 snRNP would shift the rate-limiting step to a subsequent step in the pathway.

SR proteins promote A complex formation in the absence of U1 snRNP

The binding of U2 snRNP to pre-mRNAs usually occurs after the formation of the E complex, which contains U1 snRNP, U2AF and SR proteins (Staknis and Reed, 1994). However, in the absence of U1 snRNP, the addition of a three-fold excess amount of SC35 is sufficient to form a stable U2 snRNP complex with the substrate. As would be anticipated from these observations, mutations in the 5’ splice site do not affect the rate of formation of the U2 snRNP complex in the presence of high amounts of SC35. This bypass reaction is sequence dependent, however, since it is not observed with the PIP85.A pre-mRNA. This suggests that the SR proteins bypass a U1-dependent step by promoting the binding of U2 snRNP to the branch site. SR proteins may also have a role in recruiting the U4/5/6 tri-snRNP to the U2 snRNP–substrate complex as well (R.F.
Roscigno and M.A. García-Blanco, personal communication; Tarn and Steitz, 1995). The substrate specificity of this bypass reaction may reflect the potential differences in affinities of the SR proteins for the various substrates.

Preliminary studies with chimeric pre-mRNAs containing sequences from β-globin and PIP85.A have demonstrated that the 3' portion of the β-globin RNA (comprising the branch site, polypyrimidine tract and the 3' exon), not the 5' splice site region (upstream of the branch site), confers the ability to respond to the excess SR proteins (data not shown). Thus SR proteins may act in U1 snRNP-depleted extracts by binding to a downstream sequence element and promoting the formation of A complex. The activity of SR proteins in the reconstituted reaction is reminiscent of that of exon splicing enhancer sequences. These sequence elements, which bind SR proteins, are thought to promote splicing of upstream introns by increasing the binding of U2AF to the branch site, and thus assisting in U2 snRNP recruitment (Wang et al., 1995).

The role of U1 snRNP in RNA splicing

U1 snRNP was originally shown to be required for splicing by inhibiting its activity with either a specific antibody or cleavage with RNase H (for review, see Steitz et al., 1988). Several genetic studies subsequently demonstrated that the 5' end of U1 snRNA associates with the 5' splice site by base-pairing (Zhuang and Weiner, 1986; Siliciano and Guthrie, 1988). Some mutations of the 5' splice site were suppressed by compensatory changes in U1 snRNA, but in many cases the compensatory changes did not restore splicing activity (Séraphin et al., 1988; Siliciano and Guthrie, 1988; Séraphin and Rosbash, 1990).

We now know that the same 5' splice site sequences are recognized subsequent to U1 snRNA recognition, explaining this previously observed lack of complementation. These results are consistent with the hypothesis that U1 snRNP's primary role is to
promote the assembly of U2 snRNP on the branch site, perhaps by facilitating the binding of SR proteins to the substrate RNA. The subsequent interaction of SR proteins and the 70kD protein bound to the U1 snRNP with the SR factor U2AF could increase the rate of binding of U2 snRNP (Kohtz et al., 1994; Wu and Maniatis, 1993). U1 snRNP is not, however, an important component of the spliceosome during execution of the chemistry of splicing.

In vivo, most pre-mRNAs are processed in the U1-dependent manner, but there may exist alternative pathways for splicing. U1 snRNP could have a limited role in the splicing of some pre-mRNAs, specifically those that have a high affinity for SR proteins. In the nucleus, SR proteins are localized in speckles (Fu and Maniatis, 1990; Blencowe et al., 1994). Splicing may occur within these speckles and on their periphery (Carter et al., 1993; Xing et al., 1993) where the concentration of SR proteins may be sufficiently high to promote splicing without U1 snRNP. This secondary splicing pathway would give the cell an additional way to regulate the splicing of different classes of pre-mRNAs.
MATERIALS AND METHODS

Pre-mRNA substrates
Point mutants +1 G → A, +2 U → A and +2 U → C are the previously characterized pBSAL 5, 6, and 7 plasmids respectively (Lamond et al., 1987). The other mutant β-globin pre-mRNAs were constructed by PCR-based mutagenesis, and the final constructs were sequenced. Wild-type and mutant pre-mRNAs were transcribed using T3 RNA polymerase (Pharmacia).

Preparation of snRNP-depleted nuclear extracts
HeLa nuclear extracts were depleted using the antisense affinity selection method (Blencowe and Barabino, 1995). The extracts were extensively characterized for activity and extent of depletion as described (Crispino et al., 1994).

HeLa SR proteins
SR proteins were prepared from HeLa cells and treated with micrococcal nuclease as described (Zahler at al., 1992; Crispino et al., 1994). SC35 and SRp55 were overexpressed in Hi-5 cells and purified essentially as described (Fu, 1993). These insect cell-derived proteins were treated with micrococcal nuclease in the same way as the HeLa cell preparation. The individual proteins generated single bands on both Coomassie stained SDS-polyacrylamide gels and on Western blots probed with the mAbl04 monoclonal antibody (data not shown).

Splicing Reactions
Splicing reactions were carried out under standard conditions using 20% nuclear extract (Grabowski et al., 1984). SR proteins were added at levels indicated in the figure legends. For the native gel analysis in Figure 1, a fraction of the splicing reactions was

93
loaded directly onto 4% polyacrylamide (80:1) gels which were electrophoresed in 1X TG (50mM tris-glycine, pH 8.8). The remainder of the reactions were processed and analyzed on 6% polyacrylamide/8M urea (19:1) gels in 1X TBE (89 mM tris-borate, 2 mM EDTA) to assay the extent of the splicing reactions. For the splicing complex analysis in Figure 5, the splicing reactions were treated with 0.5 mg/ml heparin for 5 minutes at 30°C prior to loading on the 4% polyacrylamide (80:1) gels. The native gels were electrophoresed either at 4°C (Figure 1B) or at room temperature (Figures 1D and 5).

_Determination of relative splicing rates_

Polyacrylamide gels were quantitated using a Molecular Dynamics PhosphorImager and ImageQuant Software version 3.22. The relative rates of the first step of splicing were determined essentially as described (Query et al., 1994). The relative amount of RNA in each band was quantitated and expressed as a percentage of the total obtained from summing the intermediates, products, and the precursor. For each band, an individual background value was determined from the region directly above that band. First step rates were determined from the linear portions of the curve and were calculated in such a way as to include all RNA species which had undergone the first step of splicing. The rates of splicing of the mutants were normalized, using parallel control reactions, to that of wild-type β-globin for each reaction type.
ACKNOWLEDGMENTS

The authors thank X.-D. Fu for supplying the baculovirus clones that express individual SR proteins SC35 or SRp55. J. Kim was instrumental in purification of the individual SR proteins from Hi-5 cell lysates, and R. Issner and Y. Qiu are thanked for excellent technical assistance. B. Blencowe, A. MacMillan, J. Pomerantz, and C. Query kindly made helpful comments on the manuscript. Many thanks also to M. Siafaca for secretarial support. This work was supported by US Public Health Service grants RO1-GM34277 and RO1-AI32486 from the National Institutes of Health and partially by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute.
REFERENCES


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


Wang, Z., H.M. Hoffmann, and P.J. Grabowski. 1995. Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* **1:** 21-35.


CHAPTER IV

U1-INDEPENDENT SPLICING IN THE ABSENCE OF EXCESS SR PROTEINS
ABSTRACT

The fushi tarazu (Ftz) pre-mRNA of Drosophila is efficiently processed in nuclear extracts that are depleted of U1 snRNP. In contrast to other precursor RNAs that require the addition of high concentrations of SR proteins to facilitate the U1-independent reaction, Ftz splices efficiently in the absence of U1 snRNP with the endogenous level of SR proteins. Sequences that are necessary for the U1 bypass pathway were localized to the 5' splice site region. These sequences were sufficient to confer U1-independent splicing when transferred into the Ad1 pre-mRNA, but not PIP85.A, suggesting a requirement for additional sequence elements in the reaction.

In a related study, chimeras between β-globin and PIP85.A were assayed for splicing under different conditions to identify sequences required for SR reconstitution of U1 snRNP-depleted extracts. One of these chimeras, PIPβG-E', spliced efficiently in the absence of U1 snRNP without the addition of SR proteins. Mutagenesis of the β-globin sequences revealed that nucleotides between the branch point and polypyrimidine tract are necessary for this reaction. Although this sequence does not resemble any in Ftz or Ad1, the RNAs share one common feature: a short polypyrimidine tract. The possible implications of this observation and potential mechanisms of U1-independent splicing are discussed.
INTRODUCTION

Pre-mRNA splicing occurs in a large multi-subunit complex termed the spliceosome which forms in a stepwise manner on the substrate. The spliceosome is composed of four snRNP particles and many non-snRNP factors, including members of the SR protein family (for review, see Moore et al., 1993). In the conventional view of this pathway, U1 snRNA, in a step that commits a substrate to splicing, base pairs to the 5' splice site and promotes the subsequent binding of the U2 snRNP with the branch site. Formation of the fully assembled spliceosome requires the addition of the U4/U5/U6 tri-snRNP; in this complex, U6 snRNA specifically binds to distal sequences within the 5' splice site that were previously associated with U1 snRNA. SR proteins participate in the splicing reaction by facilitating each step of spliceosome assembly (for review, see Fu, 1995).

U1 snRNP was originally shown to be required for splicing by inhibiting its activity with either a specific antibody or cleavage with RNase H in the presence of a complementary oligonucleotide (for review, see Steitz et al., 1988). More recent experiments in which nuclear extracts were depleted for U1 snRNP by antisense affinity depletion demonstrated that U1 snRNP is required for the earliest stages of spliceosome assembly (Barabino et al., 1990). However, splicing in the absence of U1 snRNP can be efficiently restored in some substrates by the addition of an excess of SR proteins (Crispino et al., 1994). In this U1 snRNP bypass reaction, SR proteins promote the rapid association of U2 snRNP with the branch site in a manner independent of the sequence at the 5' splice site. The entire 5' splice site consensus is recognized in a rate-limiting reaction with the binding of the U4/U5/U6 tri-snRNP (Crispino and Sharp, 1995). Thus, U1 snRNP is not required for commitment, complex assembly, or catalysis of the splicing reaction in the presence of high concentrations of SR proteins.
The activity of SR proteins in complementing a U1 snRNP-depleted reaction was dependent on the particular pre-mRNA. The β-globin substrate was the most active in this reaction, whereas an adenovirus pre-mRNA substrate, Ad1, was spliced at approximately 10% of the level observed in the mock-depleted reaction. A third substrate, PIP85.A, was not active for splicing upon addition of SR proteins to U1 snRNP-depleted reactions. This substrate specificity is consistent with the model that SR proteins which contain RNA recognition motifs bind the substrate directly and stabilize the subsequent binding of U2 snRNP.

To understand in more detail the role of SR proteins in U1-independent splicing, additional pre-mRNAs were assayed for their splicing activity. Two pre-mRNAs were identified that are efficiently processed in extracts depleted for U1 snRNP without the requirement for excess SR proteins. The first pre-mRNA was derived from the *Drosophila* fushi tarazu (*ftz*) gene, which is involved in the establishment of the segmentation pattern of the *Drosophila* embryo. Mutations in *ftz* result in embryos having half the usual number of body segments (for review, see Lawrence, 1992). The splicing of a second substrate, containing sequences derived from β-globin and PIP85.A, in the absence of U1 snRNP was dependent on nucleotides between the branch site and the polypyrimidine tract. The identification of sequences necessary for U1-independent splicing should lead to a better understanding of the mechanism of this bypass reaction and the identification of required trans-acting factors. Furthermore, the observation that certain pre-mRNAs are processed in the absence of U1 snRNP with endogenous levels of SR proteins strongly suggests that this bypass pathway may be utilized in vivo.
RESULTS

Splicing of the Ftz pre-mRNA in U1 snRNP-depleted extracts

The *Drosophila* fushi tarazu (*ftz*) gene has previously been shown to be spliced accurately and efficiently in extracts from Kc (*Drosophila*) tissue culture cells or mammalian cell extracts (Reed and Maniatis, 1985; Rio, 1988). The Ftz pre-mRNA has also been assayed for splicing in both mock and U1 snRNP-depleted extracts (J. Mermoud and A. Lamond, personal communication). Notably, the Ftz precursor RNA was spliced in the U1 snRNP-depleted extracts to nearly the same extent as in mock-depleted reactions. We have reproduced these observations, confirming that the processing of this pre-mRNA does not require U1 snRNP or excess concentrations of SR proteins, while both the β-globin and the Adl pre-mRNAs were inactive in U1 snRNP-depleted extracts (Figure 1).

Sequences around the Ftz 5' splice site are necessary but not sufficient for U1 independent splicing.

The U1-independent splicing of Ftz is reminiscent of the activity of other RNAs that were also processed in the absence of U1 snRNP, but which required the addition of high concentrations of SR proteins. Since both reactions, ΔU1 and ΔU1+SR, proceed in the absence of U1 snRNP, it is likely that the bypass mechanisms are similar and that specific sequences in the pre-mRNA define the requirement for the level of SR proteins. To investigate this sequence dependence further, synthetic pre-mRNAs containing sequences from the Adl and Ftz substrates were constructed and assayed for splicing in the different extracts. An extensive study of these chimeras determined that the 5' splice site region of the Ftz RNA was sufficient to confer U1 independent splicing to the Adl
Figure 1. Ftz pre-mRNA splicing in ΔU1 extracts. The β-globin, Ad1, and Ftz pre-mRNAs were incubated in mock or U1 snRNP-depleted extract for the indicated times at 30°. Splicing products and intermediates are indicated on the left for Ad1 and β-globin (icons preceded by a bullet refer to β-globin RNAs), while the Ftz RNAs are indicated to the right. IVS-E2, lariat intermediate; IVS- lariat product; E1-E2, ligated exons; E1, 5' exon intermediate; E1-IVS-E2, pre-mRNA.
substrate, while the 3' half, including the branch point, polypyrimidine tract and the 3' exon, was not sufficient (J. Mermoud and A. Lamond, personal communication). These same studies also demonstrated that the identity of the 5' splice site, between -2 and +6 (where +1 is the first nucleotide of the intron), was not responsible for the effect; both the Ad1 and Ftz 5' splice sites were equally active for splicing in extracts depleted of U1 snRNP.

Three chimeras provided the most information about the Ftz sequences involved in the bypass reaction. The AF58 substrate, containing 41 nucleotides of the Ftz 5' exon and the first 42 nts. of the Ftz intron (beyond the +6 position of the 5' splice site) fused to the Ad1 3' RNA, spliced in the U1 snRNP-depleted extract to nearly the same extent as in the mock-depleted reaction (Figures 2A and B). In contrast, the AdAR substrate, similar to Ad1, but with an internal deletion which shortened the intron, was inactive in the absence of U1 snRNP. The addition of SR proteins to the depleted reactions promoted the splicing of the AdAR pre-mRNA as expected, and further stimulated the U1-independent splicing of the Ftz containing substrate (Figure 2B).

The two additional chimeras, AF59 and AF60, further showed that substitution of either the Ad1 5' exon by 41 nts. of Ftz or the replacement of the 5' end of the Ad1 intron by 42 nts. of the equivalent region of the Ftz intron, respectively, conferred U1-independent splicing (Figures 2A and B). The splicing of these constructs, each containing one half of the Ftz sequences in the AF58 element, suggests that at least two sequence elements must exist. The efficiency of splicing for either of these chimeras in the U1 snRNP-depleted extract was reduced relative to that of the larger substitution, AF58, indicating that both sequence elements were required for full activity. Optimal splicing in the U1 snRNP-depleted extract was restored by the addition of excess SR proteins (Figure 2B). These 83 nucleotides of Ftz are most likely SR binding sites or sequences that bind a factor that interacts with SR proteins to facilitate splicing in the absence of U1 snRNP.
Figure 2. Sequences in the 5' region of Ftz are necessary for U1-independent splicing. 2A. Diagram of the Ad1/Ftz chimeras. These RNAs were gifts of Jacqueline Mermoud and Angus Lamond, except for AdΔR, which was subcloned by the author. The 5' splice sites are indicated. 2B. Splicing reactions of the chimeras in A. The reactions were performed for 90 min. in 20% nuclear extract supplemented with 450 ng of SR proteins or an equal volume of buffer. Splicing intermediates and products are indicated on the left.
Sequences from AF59 and AF60 were assayed for their contribution to U1 independent splicing by deletion analysis. A series of ten nucleotide deletions were made in both AF59 and AF60 by site-directed mutagenesis. None of these deletions adversely affected splicing in either mock or U1 snRNP-depleted extracts (data not shown). These Ftz sequences therefore likely contain multiple redundant elements that facilitate splicing in the absence of U1 snRNP.

To determine whether these sequences are sufficient for U1-independent splicing in other RNA contexts, chimeras were made between Ftz and PIP85.A. Various blocks of the Ftz 83 nucleotide element described above were inserted into the 5' splice site region of PIP85.A. None of the resulting constructs were active for splicing in the absence of U1 snRNP, and splicing could not be restored with additional SR proteins. When the 3' half of a Ftz/PIP85.A chimera was replaced with analogous sequences from β-globin, the U1 independent splicing process was partially restored (Crispino and Sharp, data not shown).

These data indicate that the sequences flanking the 5' splice site of the fushi tarazu pre-mRNA are necessary but not sufficient to confer U1-independent splicing. A second element, one which is present in Ad1, Ftz, and possibly β-globin, but not in PIP85.A, is required for the bypass reaction. Thus, U1-independent splicing probably requires at least two sequence components.

**β-globin contains sequences that promote U1-independent splicing**

The observation that the 3' half of β-globin facilitated U1-independent splicing of a Ftz containing pre-mRNA was reminiscent of its activity in U1 snRNP-depleted reactions supplemented with excess SR proteins. A highly purified preparation of SR proteins, or the SC35 factor alone, restored splicing to depleted extracts in a substrate specific manner, suggesting that an RNA sequence, perhaps an SR protein binding site,
Figure 3. β-globin contains sequences that promote U1-independent splicing. 3A. Schematic of PIPβG-E' which contains sequences from PIP85.A and β-globin. Seventy nucleotides surrounding the branch site of PIP85.A were replaced by sixty-eight bases from β-globin. Sequences that were altered by site-directed mutagenesis are boxed. The asterisk identifies the branch adenosine. 3B, C. Splicing reactions of the PIPβG-E' mutants were processed and electrophoresed on 15% denaturing gels. The reactions contained 20% nuclear extract supplemented with either 300 ng SC35 or an equal volume of buffer. Intermediates and products are indicated to the left. The mutants E' (64), E' (65) and E' (66) were spliced once, while the other mutants were tested several times.
A

![Diagram of genetic elements and sequences](image)

- **Xba I**
  - **PIP85A**
  - **BS**
  - **(Py)n**

- **β-Globin**
  - **+250**
  - **+295**
  - **BS**
  - **(Py)n**

- **PIPBG-E**'
  - **45 nts**
  - **BS**
  - **(Py)n**

- **62**
  - **CCC**
  - **UCUG**
  - **CUAAG**
  - **AAA**
  - **AGUA**

- **63**
  - **CUAAG**
  - **CAU**
  - **GUUC**

- **64**
  - **CAU**
  - **GUUC**
  - **UACG**

- **65**
  - **AUGC**

- **66**
  - **CUUCUUUCC**

- **67**
  - 8 nt deletion
exists in those RNAs that responded to SR proteins (Crispino et al., 1994; 1995). To identify these important sequences, substrate RNAs containing sequences from β-globin and PIP85.A were constructed and assayed for splicing. Several of these chimeras spliced in U1 snRNP-depleted extracts upon the addition of SR proteins. These results are summarized in the Appendix of this thesis.

One β-globin/PIP85.A chimera, PIPβG-E', spliced efficiently in U1 snRNP-depleted extracts even when SR proteins were omitted (Figures 3A and 3B). The PIPβG-E' RNA was constructed by removing 72 nucleotides from the PIP85.A intron and replacing these sequences with sixty-eight nucleotides of β-globin. As a result of this exchange, the new substrate has a slightly different branch site (UGCUAAC vs. UACUAAC) and a shorter polypyrimidine tract than either parental RNA (14 nts. vs. 16 nts. in β-globin and 20 nts. in PIP85.A).

It is interesting that a U1-independent substrate can be constructed by forming chimeras with two substrates that independently require U1 snRNP. The PIP85.A pre-mRNA required U1 snRNP for splicing under all conditions tested, whereas full length β-globin only spliced in the depleted extract upon addition of high concentrations of SR proteins (Crispino et al., 1994; 1995). The results are consistent with roles for two sets of sequences in splicing in the absence of U1 snRNP and suggest that some sequences in β-globin promote U1-independent splicing. The interactions between factors bound at these elements and the potential secondary and tertiary structures that these sequences form may be important in determining the requirement for U1 snRNP.

**Sequences immediately adjacent to the branch site are required for U1-independent splicing**

In both substrates described above, small blocks of RNA sequences were necessary for U1-independent splicing. To more precisely identify the downstream
sequences that are important for this splicing pathway, a panel of mutants in the β-globin portion of PIPβG-E' were constructed by site directed mutagenesis; these are summarized in Figure 3A. Four mutants were made by altering a set of nucleotides in the region between the branch site and polypyrimidine tract (mutants 62, 63, 64, and 65); these mutations were made to assay the contribution of these sequences to the bypass reaction. To understand the significance of the length of the polypyrimidine tract and the distance between the polypyrimidine tract and the branch site, two additional mutants were constructed. In the first of these, mutant 66, the eight nucleotides immediately adjacent to the branch site were changed to pyrimidines. The second mutant, 67, has these 8 nucleotides deleted. These mutations, therefore, test both the nucleotide specificity and the length of the polypyrimidine tract and its proximity to the branch site.

The activity of these mutants was assayed by splicing in mock, ΔU1, and ΔU1+SC35 extracts (Figure 3B and C). Three mutations, 64, 66, and 67, adversely affected splicing in the absence of U1 snRNP. These substrates spliced normally in mock-depleted extracts, and were fully active for splicing upon the addition of SR proteins to the ΔU1 extracts. Three other mutants, 62, 63, and 65, showed no negative effect; each was processed efficiently under all conditions.

Mutant 64, in which the three nucleotides directly adjacent to the branch site were changed to the complementary nucleotides, spliced at reduced levels in ΔU1 extracts, indicating that the identity of these bases was important. The 66 and 67 mutations, which altered the composition of the nucleotides 3' of the branch site were also inactive in ΔU1 extracts. Therefore, these eight nucleotides, or a subset thereof, deleted in E'-67 and altered in E'-66 are important for U1-independent splicing. Interestingly, these sequences were not important for splicing in the presence of high concentrations of SR proteins. That these mutations are suppressed by the addition of excess SR proteins suggests that the wild-type sequences may serve the equivalent function as SR proteins in U1 snRNP-
depleted extracts, and provides further evidence that these two splicing pathways proceed in a similar manner.

A PIP85.A construct containing only the 23 nts. of β-globin flanking the branch site, and diagrammed in Figure 3A, was not active for splicing in ΔU1 extracts, but was processed in ΔU1+SC35 reactions (Appendix, this thesis). This suggests that PIPβG-E' also requires two elements for splicing in U1 snRNP-depleted extracts when additional SR proteins are not supplied. Consistent with this, the other 45 nts. of β-globin that comprise PIPβG-E' were observed to promote splicing in ΔU1+SC35 extracts when transferred to PIP85.A, indicating that a second element may reside within these 45 nts. (Appendix, this thesis). Therefore, splicing in U1 snRNP-depleted extracts supplemented with SR proteins appears to be an intermediate step of the splicing that occurs in unsupplemented extracts.

DISCUSSION

Cis-acting sequences within an RNA can circumvent the requirement for U1 snRNP in pre-mRNA splicing. Splicing of the fushi tarazu pre-mRNA was unaffected by the removal of U1 snRNP from nuclear extracts. This U1-independent splicing was dependent on sequences within the 5' splice site region, and also on some determinant of the RNA, most likely within the branch site/ 3' splice site region. Analysis of chimeras between three substrates demonstrated that sequences of β-globin between the branch site and the 3' splice site were important for facilitating splicing in the absence of U1 snRNP. Mutations in these sequences suggest that the identity of several nucleotides is critical for this effect. Further studies of these sequences will probably be informative in determining the mechanisms by which RNAs are spliced in the absence of U1 snRNP.

Possible mechanism of U1-independent splicing
The identification of sequences that mediate splicing in the absence of U1 snRNP provides insight into the possible mechanisms of this bypass reaction. In the simplest model, these cis-acting sequences are sites of SR protein binding. Unlike U1-independent splicing reactions that occur only in the presence of excess SR proteins, splicing of the Ftz pre-mRNA is efficient with the endogenous level of SR proteins. This difference may be the result of an increased affinity for SR proteins, which therefore are required at a lower effective concentration. The mechanism may be similar to that described previously (Chapter 3; Crispino and Sharp, 1995).

Splicing in the absence of U1 snRNP appears to require a second element near the branch site. These sequences, identified in β-globin as those between the branch site and polypyrimidine tract, were shown by mutagenesis to be necessary for U1 independent splicing. These sequences could act in several ways. First, they may constitute a high affinity binding site for an SR protein. The sequence GTTCATGC, deleted from PβG-E' in mutants 66 and 67, resembles an optimal SC35 binding site, GTTCGAGTA (Tacke and Manley, 1995). The removal of this element could explain the reduced efficiency of ΔU1 splicing. This model could be tested by performing binding assays with purified SC35 and the RNA; specific binding to the wild-type sequence and not the mutants would support such a mechanism.

Second, the element may efficiently bind a different splicing factor that promotes a bypass of U1 snRNP. Several factors recognize the branch site during splicing (MacMillan et al., 1994; Gaur et al., 1995). These factors may be components of the U2 snRNP, or non-snRNP factors. The stimulation of the binding of one of these factors might facilitate complex A assembly in the absence of U1 snRNP. Factors that specifically interact with these sequences could be identified by UV cross-linking; these experiments would be especially informative when comparing the associated proteins under different extract conditions.
Alternatively, the second U1-independent element might act by virtue of not binding a nuclear factor whose activity is to repress complex assembly. Such factors might include hnRNPs and the PTB protein (for review, see Moore et al., 1993). The lack of such a non-productive interaction could allow for rapid U2AF binding and complex A formation.

An interesting possibility is suggested by the correlation between activity in U1 snRNP-depleted extracts and the length of the polypyrimidine tract \( (P_{yn}) \). The PIP85.A pre-mRNA has an extended tract of 20 nts; this RNA is not processed in the absence of U1 snRNP under any conditions. In contrast, the Ftz precursor, which has full activity in the absence of U1 snRNP, contains a short stretch of 11 pyrimidines. Similarly, PIPβG-E' which has a 13 nt. \( P_{yn} \) also spliced efficiently in \( \Delta U1 \) extracts. The Ad1 and β-globin substrates have tracts of intermediate lengths of 14 nts. and 16 nts., respectively. When the 3' halves of these RNAs were fused to the upstream Ftz element, the Ad1 fusion was the most active in \( \Delta U1 \) extracts, while the β-globin chimeras were active to lesser degrees. To test the impact of the \( P_{yn} \) length, this sequence could be altered in other RNAs, such as PIP85.A. The prediction that a shorter \( P_{yn} \) stimulates U1 independent splicing could be tested by assaying splicing in these modified RNAs.

A shorter polypyrimidine tract might facilitate U1-independent splicing by weakening the interaction of the substrate with U2AF. Although SR proteins facilitate splicing in the case of exon enhancers by increasing the binding of U2AF, the role of U2AF in splicing in U1 snRNP-depleted extracts is unknown. SR proteins might stabilize the interaction of U2AF to the substrate, or instead act by bypassing the commitment complex entirely. Interestingly, SC35 has been shown to reconstitute U2AF-depleted extracts, suggesting that SR proteins bypass E complex altogether (A. MacMillan, P. McCaw and P. Sharp, personal communication). Furthermore, this splicing reaction follows the same substrate specificity as \( \Delta U1 + SR \) splicing.
To study the role of U2AF in U1-independent splicing, several approaches to block U2AF function could be taken. Co-depletion of U2AF and U1 snRNP from nuclear extracts would provide a useful reagent to assay the role of this splicing factor in U1-independent splicing. Alternatively, U2AF antibodies, or a competitor RNA containing a U2AF binding site, could be utilized to investigate its role in the U1 bypass pathway.

Importance of ftz splicing in flies

Expression of the ftz gene is tightly controlled both temporally and spatially in the developing fly. Its expression is restricted from early blastoderm to the gastrula stage of embryonic development and maintained in only seven evenly spaced bands of cells. This expression pattern is a critical determinant in proper segmentation of the embryo. Mutations in ftz are lethal, resulting in embryos that have only half the normal number of segments. This tight regulation requires that the gene be expressed at high levels for a short period of time in specific cells. The precise means by which the cell regulates this gene’s expression is unknown.

Ftz pre-mRNA contains a single intron of 150 nucleotides, with two large exons (880 and 980 nts.). Since the RNA splices well in extracts that have been depleted of U1 snRNP, it is possible that U1 snRNP’s role in splicing this transcript in vivo is limited. Further experiments are necessary to more precisely define the important cis-acting sequences in ftz and to identify trans-acting factors that interact with these sequences. These studies may be informative in understanding the mechanism of U1-independent splicing and possibly give insight into the control of gene expression during oogenesis.

MATERIALS AND METHODS

Pre-mRNA substrates
β-globin/PIP85.A chimeras were constructed by PCR subcloning. Site-directed mutagenesis was performed with the Sculptor kit (Amersham) according to the manufacturer's instructions. All the final constructs were sequenced. Wild-type, chimeric, and mutant pre-mRNAs were transcribed using T7 or T3 RNA polymerase (Pharmacia).

**Preparation of snRNP-depleted nuclear extracts**

HeLa nuclear extracts were depleted using the antisense affinity selection method (Blencow and Barabino, 1995). The extracts were extensively characterized for activity and extent of depletion as described (Crispino et al., 1994).

**HeLa SR proteins**

SR proteins were prepared from HeLa cells and treated with micrococcal nuclease as described (Zahler et al., 1992; Crispino et al., 1994). SC35 was overexpressed in Hi-5 cells and purified essentially as described (Fu, 1993). Insect cell-derived protein was treated with micrococcal nuclease in the same way as the HeLa cell preparation. The purified protein generated a single band on both Coomassie stained SDS-polyacrylamide gels and on a Western blot probed with the mAb104 monoclonal antibody (data not shown).

**Splicing Reactions**

Splicing reactions were carried out under standard conditions using 20% nuclear extract (Grabowski et al., 1984). SR proteins were added at levels indicated in the figure legends. The reactions were processed and analyzed on 8% or 15% polyacrylamide/8M urea (19:1) gels in 1X TBE (89 mM tris-borate, 2 mM EDTA) to assay the extent of the splicing reactions.
ACKNOWLEDGMENTS

Many thanks to Jacqueline Mermoud and Angus Lamond for providing us with the Ftz pre-mRNA, as well the chimeras AF58, AF59, and AF60, and for communicating their results. Additional thanks to B. Blencowe, L. Lim, A. MacMillan, and J. Pomerantz for reading of this chapter, and to other members of the Sharp lab for their support.

REFERENCES


APPENDIX

MAPPING THE SEQUENCES IMPORTANT FOR SR COMPLEMENTATION OF U1-snRNP-DEPLETED EXTRACTS
RESULTS

As described in the preceding chapters, the efficiency with which SR proteins complemented a U1 snRNP-depleted extract was dependent on the particular pre-mRNA. β-globin RNA was processed to the same extent in ΔU1+SR reactions as in mock-depleted extracts, whereas the PIP85.A substrate was not active for splicing under these reconstituted conditions. To define the sequences in β-globin that facilitated splicing in the absence of U1 snRNP with high concentrations of SR proteins, several sets of chimeric substrates were constructed using sequences from both RNAs. The design of these constructs, as well a summary of their splicing activity in ΔU1+SR reactions, is provided in this section. Although the results do not precisely identify a sequence that promotes U1-independent splicing in the presence of elevated levels of SR proteins, the data strongly suggest that multiple elements exist. These positive sequences, found exclusively within the intron of β-globin, may act by binding SR proteins which in turn promote the binding of U2 snRNP, in a manner analogous to exon enhancers. It is likely that several elements, both positive and negative, participate in dictating the requirements for U1 snRNP.

Several sequences within the RNAs were found to influence SR reconstitution of U1 depleted extracts. One positive element was localized to the 5' portion of the β-globin intron, between +50 and +100 nucleotides (+1 being the first nucleotide of the intron). A second sequence, 23 nucleotides between the branch site and the polypyrrimidine tract of β-globin, was sufficient to confer SR responsiveness to PIP85.A. Furthermore, 45 nucleotides immediately upstream of the branch site in β-globin also restored splicing to PIP85.A in ΔU1+SR reactions. This final result could also be explained by the removal of a putative negative element from PIP85.A that may mandate the participation of U1 snRNP in the splicing process. These findings and their implications are summarized below.
A positive element exists between +50 and +100 of the β-globin intron

The first chimeras tested in the different reactions are diagrammed in Figure 1A. In this set, the β-globin 5' exon with increasing lengths of intron sequence were fused to PIP85.A at its Xba I site; thus, the branch site, polypyrimidine tract (Py₅₇) and the 3' exon were derived from PIP85.A. The first construct, βGPIP 50, containing the first 50 nucleotides of the β-globin intron, was not spliced in the absence of U1 snRNP, although it was processed in the mock-depleted extract (Figure 1B). For the remaining chimeras, with 100, 150, 200, or 250 nucleotides of the β-globin intron, SR proteins restored splicing to the U1 snRNP-depleted extract. Although the extent of reconstituted splicing was not equal to that observed in the mock-depleted extracts, there was a clear enhancement of splicing. Therefore, β-globin intron sequences between +50 and +100 provided a signal for SR activity. Analysis of this region identified two purine-rich sequences: GGAGGGGCGAAAAG at +61, and AGAATGGGAAGA at +91. These purine-rich elements may facilitate splicing by binding SR proteins and stabilizing the interactions between splicing factors and the substrate.

Analysis of the 3' half of β-globin

To test the contribution of the 3' portion of β-globin to the SR reconstituted splicing reaction, the 5' half of PIP85.A (up to the Xba I site) was fused to β-globin at position +250 in the intron (Figure 2A). This chimera, PIPβG, contains the branch site, polypyrimidine tract, and 3' exon of β-globin. The RNA was fully active in the AU1+SC35 reactions, splicing to the same extent as in mock-depleted extracts (Figure 2B). Since this chimera spliced more efficiently in the reconstituted reactions than the βGPIP series described above, subsequent experiments were focused on characterizing these 3' sequences of β-globin.

Residues in PIPβG were mutated to define important sequences derived from the β-globin RNA. The first three mutants, Δ51, Δ52, and Δ53, each contained 15 nt.
Figure 1. Splicing of βGPIP chimeras. 1A. Chimeras between β-globin and PIP85.A were constructed by PCR. An increasing amount of β-globin intron sequence was fused to PIP85.A at its Xba I site. The shaded boxes and wavy lines are sequences derived from β-globin. BS, branch site; (Py)n, polypyrimidine tract. 1B. Splicing assays of the chimeras diagrammed in A. Splicing reactions containing 20% nuclear extract and either 450 ng SR proteins, or an equal volume of buffer were incubated for 90 minutes at 30°C. The reactions were treated with phenol and ethanol precipitated prior to electrophoresis on this 8% denaturing gel. The intermediates and products are indicated on the left: IVS-E2, lariat intermediate; IVS, lariat product; E1-E2, ligated exons; E1, 5′ exon intermediate; E1-IVS-E2, pre-mRNA. The addition of SR proteins to the U1 depleted extract promoted splicing to a cryptic 5′ splice site when the β-globin intron contribution was extended from +100 to +150 nucleotides. The mobilities of the intermediates and products of the reaction at the cryptic site are indicative of the activation of a proximal splice site.
deletions in the β-globin region between +250 and +295 (Figure 2A). These mutants, which together delete the 45 nucleotides of β-globin immediately 5' of the branch site, had no effect on splicing in either mock, ΔU1, or ΔU1+SC35 reactions (Figure 2B). These results suggest that none of these sequences play an important role in the SR reconstitution of U1 snRNP-depleted extracts in this substrate. In a fourth mutant, four nucleotides immediately downstream of the branch adenosine were scrambled to assay their significance (Figure 2A). These positions were targeted because the sequence resembles an optimal SC35 binding site (Tacke and Manley, 1995), and the disruption of this element was predicted to specifically inhibit ΔU1+SC35 splicing. However, this mutation also had no effect on the extent of splicing in either extract (data not shown).

23 nucleotides from β-globin are sufficient to confer ΔU1+SC35 splicing to PIP85.A

To more precisely map the location of relevant sequences, several additional chimeras were made (Figure 3A). First, to assay the role of the 3' exon, two chimeras in which portions of the PIP85.A 3' exon were replaced by analogous sequences in the β-globin 3' exon were synthesized (PIPβG-A and PIPβG-B). Three other RNAs were constructed in parallel to test the remaining β-globin contribution. These three, PIPβG-C, PIPβG-D, and PIPβG-E, contained progressively less β-globin sequence fused to PIP85.A at its Xho I site (Figure 3A). These RNAs did not contain the 45 nts. deleted from PIPβG in RNAs Δ51, Δ52, and Δ53 (β-globin +250 through +295). Instead, PIP85.A sequences between the Xba I and Xho I site, not present in PIPβG, were included to maintain a constant intron size.

The two exon swaps, PIPβG-A and PIPβG-B were not processed in ΔU1+SR reactions over the background level of PIP85.A splicing (Figure 2B). The three other chimeras, however, were efficiently processed in U1 snRNP-depleted extracts supplemented with SR proteins. One of these, E, contains only 23 nts. of β-globin in
Figure 2. PIPβG chimera spliced in ΔU1+SC35. 2A. Schematic of the PIPβG chimera. Sequences in β-globin between +250 in the intron and +50 of the 3' exon were amplified by PCR to incorporate an Xba I site at the 5' end and a Hind III site at the 3' end. The resulting DNA was inserted into PIP85.A at its Xba I/ Hind III sites to produce the chimera. The four mutations were introduced by site-directed mutagenesis. The shaded boxes and wavy lines are sequences derived from β-globin. BS, branch site; (Py)n, polypyrimidine tract. The asterisk identifies the branch adenosine. 2B. Splicing of the PIPβG RNA and three deletion mutants. The splicing reactions were performed as in Figure 1 except that 300 ng SC35 was added instead of the HeLa SR proteins. SC35 and HeLa SR proteins gave identical results in the splicing assays.
Figure 3. Splicing of PIPβG chimeras. 3A. PIP85.A RNAs containing various insertions of β-globin are diagrammed. The RNAs were constructed by PCR. The shaded boxes and wavy lines are sequences derived from β-globin. BS, branch site; (Py)n, polypyrimidine tract. 3B. Splicing of the PIPβG chimeras under the three different extract conditions. As in Figure 2, 300 ng SC35 or an equal volume of buffer were added to the extracts. Splicing reactions and icons are as described in Figure 1. The splicing of the PIPβG-C chimera in the ΔU1+SC35 reaction in this figure is not representative of its activity: the RNA was degraded during processing, so the efficient reconstitution is not as evident.
PIP85.A. This short sequence was sufficient to promote U1-independent splicing by SR proteins, or SC35. Interestingly, the 23 nt. substitution encompasses the branch site, which only deviates from consensus and PIP85.A in a single position (UGCUAAC vs. UACUAAC). The substitution is not likely to significantly alter the binding of U2 snRNA, because the base pairing potential is only slightly affected. These sequences might facilitate splicing in several ways.

First, these sequences may be SR binding sites, and the bypass reaction is facilitated by SR proteins bound to the substrate. SRs then promote the binding of U2 snRNP to the branch site, and possibly assist in recruitment of the U4/U5/U6 tri-snRNP. Alternatively, the SR proteins may bind the pre-mRNA non-specifically, while these sequences facilitate complex A formation by interacting with other splicing factors, possibly proteins of the U2 snRNP. Experiments to determine the factors that specifically associate with these sequences, such as UV cross-linking or mobility shift assays using purified splicing factors, would be informative in understanding the mechanism by which these sequences promote splicing in the absence of U1 snRNP.

Further chimeras suggest that additional elements exist

The 45 nts. that lie between +250 and +295 of the β-globin intron were sufficient to confer ΔU1+SC35 splicing to PIP85.A when substituted for the PIP85.A sequences between Xba I and Xho I (PIP70, Figures 4A and B). This experiment provides evidence that these β-globin sequences actually contain a positive signal for SR dependent splicing in U1 snRNP-depleted extracts. This result contrasts with that obtained in the individual deletions Δ51, Δ52, and Δ53 (Figure 2), where the removal of these same bases had no effect on splicing. The simplest explanation is that these sequences mediate a positive effect on splicing in ΔU1+SC35 reactions, and that PIPβG
Figure 4. Additional chimeras of β-globin and PIP85.A RNAs. 4A. The structure of chimeras PIP70 and PIP71 are diagrammed. The shaded boxes and wavy lines are sequences derived from β-globin. BS, branch site; (Py)n, polypyrimidine tract. 4B. Splicing of the PIP70 RNA under the different reaction conditions. Splicing reactions containing 20% nuclear extract and either 300 ng SC35, or an equal volume of buffer were incubated for 90 minutes at 30°. The reactions were treated with phenol and ethanol precipitated prior to electrophoresis on this 15% denaturing gel. The intermediates and products are indicated on the left: IVS-E2, lariat intermediate; IVS, lariat product; E1-E2, ligated exons; E1, 5' exon intermediate; E1-IVS-E2, pre-mRNA.
contains multiple positive elements. Such redundancy would complicate the deletion analysis.

That these 45 nts. of β-globin contain a positive element is further supported by the splicing patterns of other chimeras. PIPβG-C, PIPβG-D, and PIPβG-E were modified by the removal of the PIP85.A Xba I-Xho I and a substitution by the 45 nts. of β-globin (+250 through +295). The resulting RNAs, PIPβG C', D', and E', consistently spliced better in the ΔU1+SR reactions than the original chimeras (data not shown). In addition, the PIPβG-E' RNA was now efficiently spliced in the U1 snRNP-depleted extracts without additional SR proteins. The data support the existence of two positive elements in the E' chimera; both appear necessary for efficient U1-independent splicing in the absence of exogenously added SR proteins. This result is further described in Chapter 4.

**Does PIP85.A contain a negative element?**

An alternative explanation for the above splicing patterns is that PIP85.A contains a negative element in the region between Xba I and Xho I. Thus, the removal of PIP85.A between Xba I and Xho I in PIP70, and not the inclusion of the β-globin sequences might be the critical determinant. We were led to consider this possibility because the individual deletions of the 45 nts. from PIPβG, Δ51, Δ52, and Δ53, did not adversely affect splicing. This negative model proposes that SR reconstitution of U1 snRNP-depleted extracts would occur with most pre-mRNAs. PIP85.A, however, is not active because it contains an element that prevents splicing in the absence of U1 snRNP.

A deletion from Xba I to Xho I in PIP85.A did not rescue splicing activity in the ΔU1+SC35 reaction (data not shown). Furthermore, a substrate in which the PIP85.A contribution to PIPβG was extended to include the Xba I through Xho I sequence (PIP 71; Figure 4A) was spliced as efficiently as PIPβG in each reaction condition (data not shown). Thus, the evidence favors the existence of multiple positive elements being
responsible for pre-mRNA splicing in U1 snRNP-depleted extracts supplemented with SR proteins.

**SUMMARY**

It is likely that splicing in U1 snRNP-depleted extracts requires multiple sequence elements. For those RNAs that spliced in these extracts with the endogenous level of SR proteins, Ftz and PIP\(\beta\)G-E', multiple positive elements appear necessary. Furthermore, the presence of other sequences can prevent this bypass splicing reaction, because full length \(\beta\)-globin cannot splice in AU1 extracts without additional SR proteins. For those reactions in U1 snRNP-depleted extracts that are supplemented by excess SR proteins, a single element may be sufficient. Further experiments to better define the sequences that promote splicing in the absence of U1 snRNP and to characterize the factors that interact with these sequences to facilitate this process will provide valuable information about both constitutive and alternative mechanisms of splicing.

**MATERIALS AND METHODS**

*Pre-mRNA substrates*

\(\beta\)-globin/PIP85.A chimeras were constructed by PCR subcloning. Site-directed mutagenesis was performed with the Sculptor kit (Amersham) according to the manufacturer's instructions. All the final constructs were sequenced. Wild-type, chimeric, and mutant pre-mRNAs were transcribed using T7 or T3 RNA polymerase (Pharmacia).

*Preparation of snRNP-depleted nuclear extracts*
HeLa nuclear extracts were depleted using the antisense affinity selection method (Blencowe and Barabino, 1995). The extracts were extensively characterized for activity and extent of depletion as described (Crispino et al., 1994).

**HeLa SR proteins**

SR proteins were prepared from HeLa cells and treated with micrococcal nuclease as described (Zahler et al., 1992; Crispino et al., 1994). SC35 was overexpressed in Hi-5 cells and purified essentially as described (Fu, 1993). Insect cell-derived protein was treated with micrococcal nuclease in the same way as the HeLa cell preparation. The purified protein generated a single band on both Coomassie stained SDS-polyacrylamide gels and on a Western blot probed with the mAAb104 monoclonal antibody (data not shown).

**Splicing Reactions**

Splicing reactions were carried out under standard conditions using 20% nuclear extract (Grabowski et al., 1984). SR proteins were added at levels indicated in the figure legends. The reactions were processed and analyzed on 8% or 15% polyacrylamide/8M urea (19:1) gels in 1X TBE (89 mM tris-borate, 2 mM EDTA) to assay the extent of the splicing reactions.

**ACKNOWLEDGMENTS**

Many thanks to B. Blencowe, L. Lim, A. MacMillan, and J. Pomerantz for assisting in the preparation of this chapter and to the other members of the Sharp lab for providing helpful advice.
REFERENCES


BIOGRAPHICAL NOTE

Education

1990- present  MASSACHUSETTS INSTITUTE OF TECHNOLOGY
                PhD. in Biology, November 1995.

1986-1990  WASHINGTON UNIVERSITY, St. Louis, MO.
                Bachelor of Science, Magna Cum Laude, May 1990.
                GPA 3.62 on a 4.0 scale. Phi Beta Kappa
                President of fraternity house.

Positions

1990-present  Graduate Research Fellow, MIT. Thesis under Professor Phillip
                A. Sharp on "The Role of U1 snRNP in pre-mRNA Splicing."
                Characterizing the ability of several mammalian proteins to
                substitute for the U1 snRNP factor in pre-mRNA splicing
                reactions. Also investigating the mechanism of this splicing
                pathway and its potential role in vivo.

1993  Teaching Assistant, MIT. Participated in a team of three TAs
                with Professor Anthony Sinskey to design and teach an
                advanced laboratory course in bacterial genetics to MIT
                undergraduates, while conducting thesis research. Project
                involved cloning genes in the methionine biosynthetic pathway
                in E. coli.

                Performed independent research for an honors thesis in the field
                of plant biology under the supervision of Professor Roger
                Beachy. Analyzed the role of a class of soybean DNA binding
                proteins in plant development.

1988  National Science Foundation Summer Research Fellow,
                Columbia University, NY. Investigated the interaction of novel
                organometallic compounds with DNA under the supervision of
                Professor Nicholas Turro.

Research Awards

1990-1993  NIH Training Grant
1989  Sigma Xi Research Grant
1988  NSF Research Experience for Undergraduates Summer Fellow

Publications

Crispino, J.D. and Sharp P.A., "A U6 snRNA:pre-mRNA interaction can be
rate-limiting for U1-independent splicing," *Genes and Development* 9: 2314-
2323.