The EMBO Journal vol.15 no.6 pp.1392–1402, 1996

Three recognition events at the branch-site adenine

Charles C.Query, Scott A.Strobel[1,2] and Phillip A.Sharp[3]

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139-4307 and [1]Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO 80309-0215, USA

[2]Present address: Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

[3]Corresponding author

An adenosine at the branch site, the nucleophile for the first transesterification step of splicing, is nearly invariant in mammalian pre-mRNA introns. The chemical groups on the adenine base were varied systematically and assayed for formation of early spliceosome complexes and execution of the first and second steps of splicing. Recognition of constituents of the adenine is critical in formation of a U2 snRNP-containing complex on a minimal branch-site oligonucleotide. Furthermore, the efficiencies of the first and second chemical steps have different dependencies on the functional groups of the adenine. In total, the chemical groups on the adenine base at the branch site are differentially recognized during at least three different processes in the splicing of pre-mRNA. Moreover, a protein, p14, interacts with the adenine in a base-specific fashion and may mediate early recognition of this base.

Keywords: adenine analogs/nucleophile/RNA splicing/site-specific modifications

Introduction

The removal of introns from precursors to messenger RNA molecules (pre-mRNA, E1–IVS–E2) involves two transesterification reactions. In the first, the 2′-OH of an adenosine residue (the branch site) is the nucleophile that attacks the 5′ splice site phosphate, resulting in a free 5′ exon (E1) and a lariat form of the intron–3′ exon (IVS–E2). In the second reaction, the 3′-OH of the 5′ exon attacks the 3′ splice site, resulting in release of the lariat intron (IVS) and ligation of the two exons (E1–E2) (reviewed in Moore *et al.*, 1993; Madhani and Guthrie, 1994). Clearly, the branch-site adenosine is involved chemically in the first step and not in the second step; nevertheless, the identity of its base as adenine is important for both steps. In particular, an adenine to guanine base change imparts a strong block to step two, as well as a significant decrease in step one (see Hornig *et al.*, 1986; Freyer *et al.*, 1987; Query *et al.*, 1994). This suggests that the branch is important for interactions specific to the second step active site. Consistent with this interpretation, the second step requires an interaction between the 5′ splice site G_{+1} (the first nucleotide of the intron, now attached to the branched adenosine) and the 3′ splice site G_{-1} (the last nucleotide of the intron), placing the branch structure immediately adjacent to the 3′ splice site (Parker and Siliciano, 1992; Scadden and Smith, 1995). Thus, interactions at and near to the branch site are critical for both spliceosomal assembly and the two transesterification reactions. In this study, the functional groups responsible for branch-site identity were directly tested by incorporation of site-specific base modifications at the branch site of a pre-mRNA.

The two splicing reactions are catalyzed by the spliceosome, a dynamic 50–60S complex comprised of the small nuclear RNAs (snRNAs) U1, U2, U5 and U4/6, as well as protein components (for review, see Lamm and Lamond, 1993; Moore *et al.*, 1993; Madhani and Guthrie, 1994; Nilsen, 1994). Commitment of a pre-mRNA to the splicing pathway involves formation of the E (early) or commitment complex, which contains U1 snRNP as well as non-snRNP protein factors, including the factor $U2AF^{65}$ bound near the branch region (Séraphin and Rosbash, 1989; Michaud and Reed, 1991; Jamison and Garcia-Blanco, 1992). Native gel analysis allows the separation of three later splicing complexes, termed A, B and C, from the non-specific H complex (Konarska and Sharp, 1986, 1987). Complex A is generated by the stable binding of U2 snRNP to the branch region of the pre-mRNA. A larger complex, B, is formed by association of U4–U5–U6 tri-snRNP with complex A. Complex C, which follows B, contains splicing intermediates, but does not contain stably associated U1 or U4 snRNPs.

Recognition of the branch region is required early in splicing complex assembly. The branch sequence 5′-UACUAAC-3′ (the underlined adenosine is the nucleophile for the first step) is highly conserved in yeast (reviewed in Rymond and Rosbash, 1992); in mammals, the consensus sequence is 5′-$^{C}_{U}U^{A}_{G}$AC-3′, although the yeast sequence remains optimal (Reed and Maniatis, 1988; Zhuang *et al.*, 1989). U2 snRNP recognizes the intron in part through U2 snRNA–branch region base-pairing (Parker *et al.*, 1987; Wu and Manley, 1989; Zhuang and Weiner, 1989) and the first step nucleophile is selected in part by virtue of being bulged from this duplex (Query *et al.*, 1994). A number of protein factors recognize the pre-mRNA in the vicinity of the branch region. $U2AF^{65}$ binds specifically to polypyrimidine tracts (Ruskin *et al.*, 1988; Zamore *et al.*, 1992) and is required for stable branch region–U2 snRNA interactions. Early branch-site recognition in yeast is mediated in a manner dependent on U1 snRNP by an as yet unidentified non-snRNP splicing factor, one component of which is MUD2, a potential homolog of $U2AF^{65}$ (Ruby and Abelson, 1988; Séraphin and Rosbash, 1991; Abovich *et al.*, 1994). Mammalian factors SF3a and SF3b, in addition to $U2AF^{65}$

and U1 snRNP, are required for the stable association of U2 snRNP with the pre-mRNA branch region (Behrens *et al.*, 1993; Brosi *et al.*, 1993a,b; Krämer, 1995). Some of these U2 snRNP-associated proteins can be cross-linked to the region of the branch sequence in complex A and in spliceosomes (Staknis and Reed, 1994); for example, one of these, spliceosome-associated protein (SAP) 49, is a component of SF3b and cross-links to the pre-mRNA 5′ to the branch region (Champion-Arnaud and Reed, 1994). Yeast protein PRP16p has been proposed to be part of a proofreading/discard pathway which examines the branched nucleotide; mutant prp16 alleles allow the progression of certain non-adenine branches to continue through the second step of splicing (Couto *et al.*, 1987; Burgess and Guthrie, 1993). Recently, using a highly active photoreagent placed site-specifically on the branch nucleotide to probe within 15 Å, six proteins were detected that rearranged during spliceosomal assembly: an early p80 species; p14, p35 and p150 proteins in complex A; and p14, p70 and p220 species in the spliceosome (MacMillan *et al.*, 1994). In addition, using a thio-uridine cross-linker placed 3′ to the branch sequence, U2AF^{65} could be detected during early complex formation and a novel protein, p28, was detected after the first chemical step (Gaur *et al.*, 1995). However, none of these proteins, as yet, have been shown to impart specific branch-site adenine recognition.

In the present study, the role of functional constituents of the adenine base in branch-site identity during mammalian spliceosome assembly, as well as in the two chemical steps of splicing, was investigated by incorporation of different modified bases at this position. Further, direct UV cross-linking with these modified RNAs identified one of the above factors, p14, as intimately associated with the adenine.

Results

Splicing of branch-site base-modified E1–IVS–E2 RNAs

Previous studies have shown that an adenine to guanine mutation at the branch site is inhibitory for both of the chemical steps of splicing, suggesting that there is specific recognition of the adenine base and/or that some group on the guanine is inhibitory (Hornig *et al.*, 1986; Freyer *et al.*, 1987; Query *et al.*, 1994). To evaluate the contributions of different functional groups on the branch-site base, pre-mRNAs were constructed containing at the preferred branch-site position either 2,6-diaminopurine (also known as 2-amino-adenine, which has an exocyclic C6-NH_2 group as does adenine and an exocyclic C2-NH_2 group as does guanine), purine (the base of nebularine, which has no groups at positions C2 and C6) or hypoxanthine (the base of inosine, which has a C6-oxo group and a protonated N1-H as does guanine) (see Figure 1A). The effects of these modifications on splicing and the formation of splicing complexes were compared with pre-mRNAs containing adenine, guanine, cytosine or uracil at the branch site (Figure 2). Comparison of adenine with purine allows direct evaluation of the C6-NH_2 group; comparisons both of guanine with hypoxanthine and of 2,6-diaminopurine with adenine allow evaluation of the C2-NH_2 group; and comparisons of hypoxanthine with purine and of guanine with 2,6-diaminopurine allow evaluation of constituents at positions N1 and C6 together.

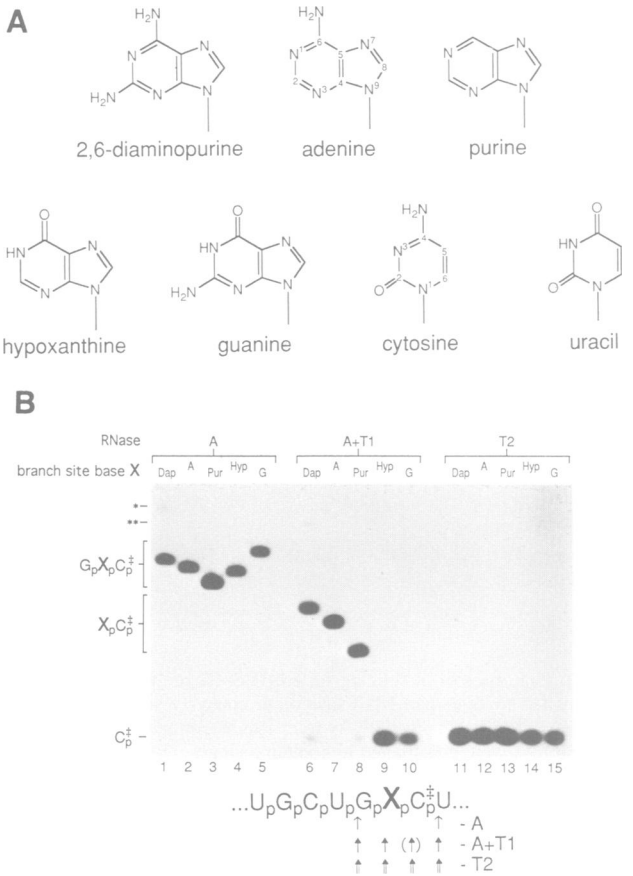

Fig. 1. (**A**) Structures of purine base analogs compared with adenine, guanine, cytosine and uracil bases. Ring positions of adenine and cytosine are numbered to facilitate comparisons. (**B**) Nuclease analysis of branch-site modifications. A 25% polyacrylamide gel of site-specifically labeled E1–IVS–E2 RNAs digested with ribonucleases A (lanes 1–5), A+T1 (lanes 6–10) or T2 (lanes 11–15). E1–IVS–E2 pre-mRNAs were constructed to contain a single ^{32}P-labeled phosphate (‡) as indicated below the lanes. The modified base at the preferred branch-site position (X) is indicated. Below the panel is shown the branch-site region with the relevant cleavage sites indicated for RNase A (↑, upper row of arrows), RNases A+T1 (⇑, middle row of arrows) and RNase T2 (⇑, lower row of arrows); parentheses indicate an RNase T1 cleavage site if base X is hypoxanthine (Hyp) or guanine (G), but not 2,6-diaminopurine (Dap), adenine (A) or purine (Pur). Structures of the predominant oligonucleotides are indicated. Two minor products are indicated by asterisks: * in lane 1 and ** in lane 6 are specific to 2,6-diaminopurine-containing oligonucleotides and may represent a small percentage (<5%) of partially protected bases. A small amount of $C_p^‡$ in lanes 6 and 8 is due to a very slow rate of cleavage by RNase T1 after 2,6-diaminopurine and purine nucleotides, respectively. Principal bands are purposefully overexposed to demonstrate the absence of detectable contaminants. There was no contaminating adenine in the branch-site modified RNAs (see text) and all the modifications are cleaved by RNase T2 (lanes 11, 13 and 14).

Pre-mRNAs containing the modified bases did not contain detectable contamination with adenine at the branch position as determined by the electrophoretic mobilities of the resulting tri- and dinucleotides following digestions with RNases (Figure 1B). 2,6-Diaminopurine- and purine-containing oligonucleotides clearly migrated differently from adenine-containing oligonucleotides (compare lanes 1 and 3 with 2; 6 and 8 with 7). Hypoxanthine-containing RNA trinucleotide migrated differently

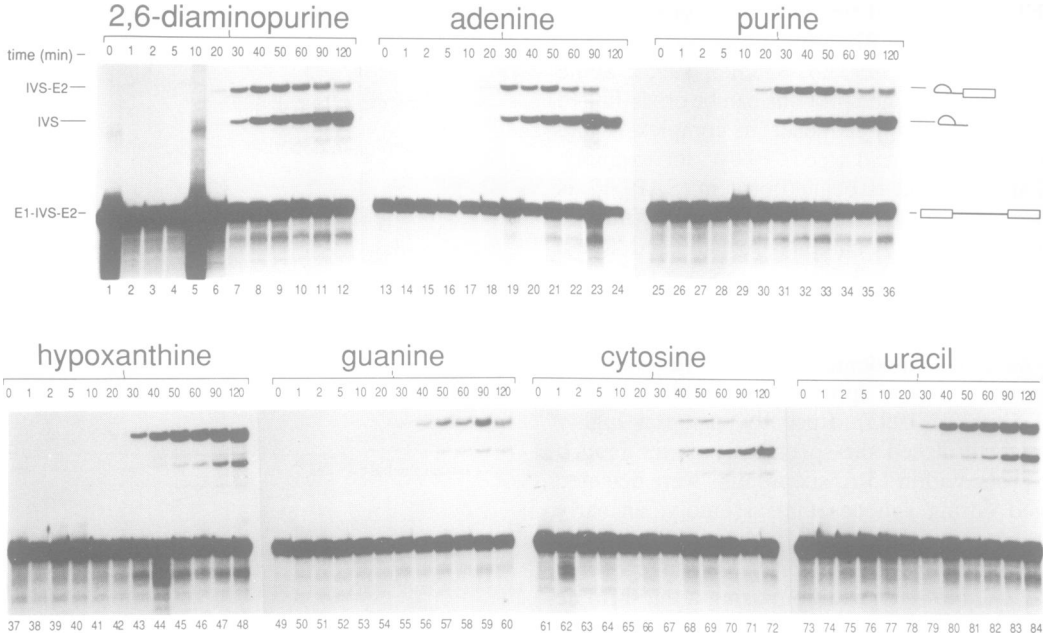

Fig. 2. Time courses of splicing for E1–IVS–E2 pre-mRNAs containing base modifications at the branch-site nucleotide. Pre-mRNAs were constructed by three-part RNA ligation to contain a single modification at position X within the strong branch sequence 5′-...UGCUGXC...-3′. Labeled E1–IVS–E2 pre-mRNAs were incubated under splicing conditions for the times indicated and then separated on a 15% polyacrylamide gel. Pre-mRNAs were constructed to contain a single ^{32}P label one nucleotide 3′ of the branch-site position to allow subsequent analysis of branch-site usage (see Figure 3 and Table I); thus, only intron-containing species are visualized. (IVS–E2) the lariat form of the intron–3′ exon intermediate; (IVS) the excised lariat intron; (E1–IVS–E2) the pre-mRNA.

from guanine (compare lanes 4 and 5) and, since it was completely cleaved by RNase T1, it contained no detectable adenine (compare lanes 9 and 7). Similarly, cytosine- and uracil-containing pre-mRNAs did not contain any contaminating purine nucleotides at the branch position, since they were completely cleaved by RNase A (data not shown and Figure 3, lanes 31 and 37).

The effects of the base modifications on branch recognition and utilization were determined by examining the splicing of E1–IVS–E2 pre-mRNAs containing a single modified base. The 2,6-diaminopurine and purine branch sites were spliced at similar rates to adenine, with a slight decrease in the rate of step one but little or no effect on step two (Figure 2, compare lanes 1–12, 13–24 and 25–36). In contrast, the hypoxanthine branch site was spliced similarly to guanine and uracil—with a decrease in the rate of step one and a strong inhibition of step two (compare lanes 37–48, 49–60 and 73–84). The cytosine branch site had a reduced rate for step one, but efficiently underwent step two (lanes 61–72).

Because either of two adjacent positions in the branch sequence can be utilized as the nucleophile in the first step (Query *et al.*, 1994), the splicing products were analyzed to determine the exact usage of the modified nucleotides (Figure 3). The pre-mRNAs contained a single ^{32}P label one nucleotide 3′ to the modified position (as indicated by ‡ under each panel in Figure 3; see also Materials and methods). Thus, digestion of E1–IVS–E2 pre-mRNAs with RNase T2, which cleaves 3′ to any ribonucleotide (Figure 1B, lanes 11–15; Uchida and Egami, 1971), generates $C_p^‡$, as expected (Figure 3, lanes 4, 10, 16, 22, 28, 34 and 40). Since RNase T2 does not cleave within branched trinucleotides (Wallace and Edmonds, 1983), digestion of the IVS–E2 or IVS lariat RNAs could produce two possible species—a labeled branched trinucleotide, $X_pC_p^‡$ (where the superscripted residues indicate the 2′–5′ linkage), only if the modified position X was used as the first step nucleophile; or a labeled mononucleotide, $C_p^‡$, if the 5′-adjacent guanosine or if some other branch region was activated as the first step nucleophile. Use of the modified position X as the first step nucleophile was assayed by the relative amounts of $X_pC_p^‡$ and $C_p^‡$ in IVS–E2 and IVS RNAs (Figure 3, lanes 5–6, 11–12, 17–18, 23–24, 29–30, 35–36 and 41–42; see Table I). Activation of other branch regions was also monitored by digestion with RNase A, which cleaves 3′ to any pyrimidine (Blackburn and Moore, 1982), resulting in a labeled branched species for either of the adjacent positions $G_pX_pC_p^‡$ or $G_pX_pC_p^‡$ where X is any purine, but resulting in $G_pX_pC_p^‡$ trinucleotide if branch formation were at some other cryptic site (Figure 3, lanes 2–3, 8–9, 14–15, 20–21 and 26–27).

Taken together, the RNase A and T2 digests revealed that 2,6-diaminopurine and purine were selected in branch formation relative to other sites nearly as efficiently as adenine—with the ratio of usage to all other sites being ~30:1 compared with 40:1 for adenine. Hypoxanthine was selected less in branch usage, ~2.5:1; and guanine, as expected, was selected the least at ~1:1. Cytosine and uracil were both used moderately, at 10:1 and 5:1, respectively (Table I).

The data for utilization of branch position were used to calculate the rates of the first and second reactions of

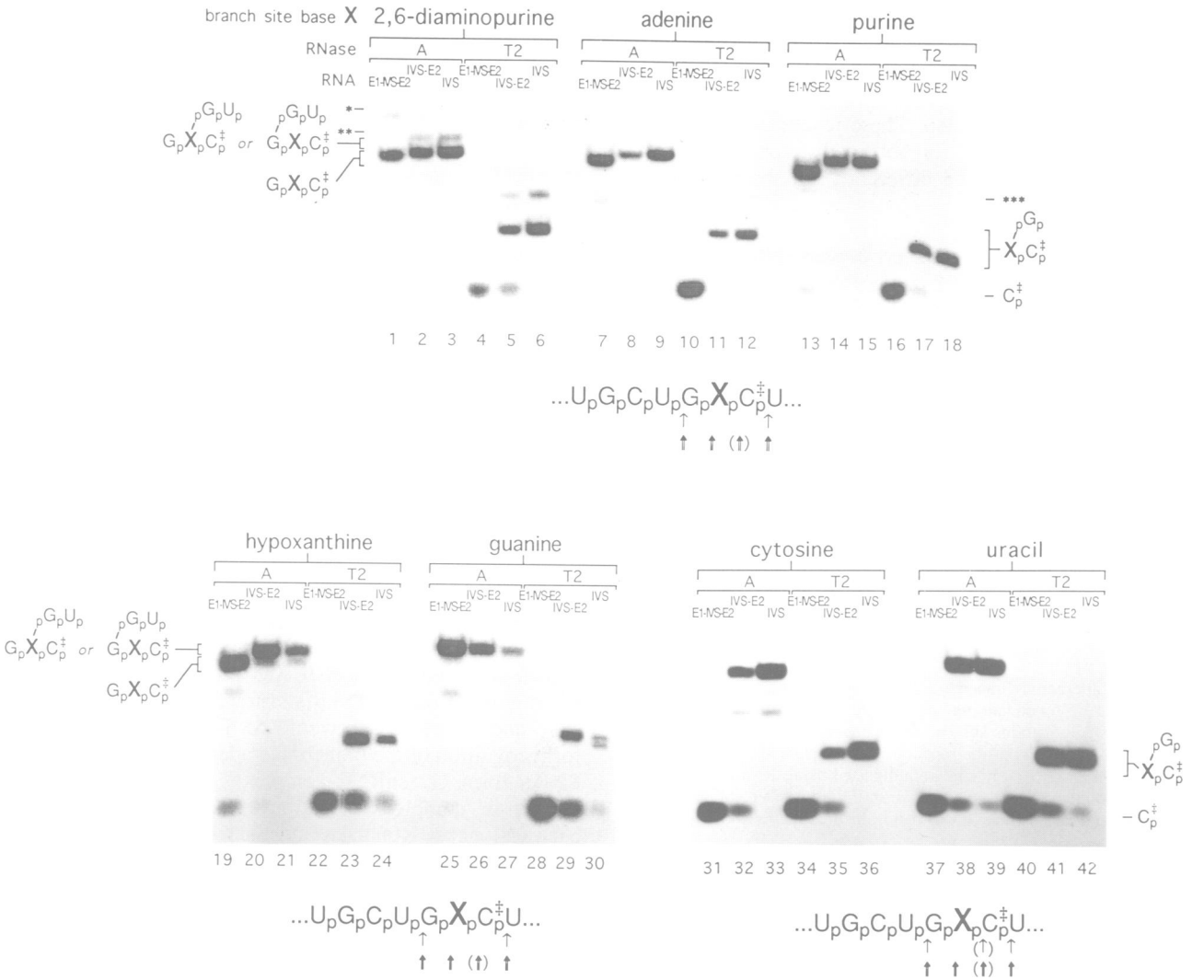

Fig. 3. Nuclease analysis of branch sites. A 25% polyacrylamide gel of site-specifically labeled splicing intermediates and products digested with RNase A or T2. E1–IVS–E2 pre-mRNAs were constructed to contain a single ^{32}P-labeled phosphate (±) as indicated below the lanes. The base at the preferred branch-site position (X) is indicated above each set. After incubation of E1–IVS–E2 pre-mRNAs under splicing conditions for 90 min, all species were purified on a 15% polyacrylamide gel. Subsequently, the E1–IVS–E2, IVS–E2 and IVS RNAs were digested with RNase A or T2. Structures of the predominant oligonucleotides are indicated (RNase A products at left, RNase T2 products at right). Several minor products are indicated by asterisks: * in lane 1, ** in lanes 2 and 3, and *** in lanes 5 and 6 are specific to 2,6-diaminopurine-containing oligonucleotides as in Figure 1B. Below each panel set is shown the branch-site region with the relevant cleavage sites indicated for RNase A (↑, upper row of arrows) and for RNase T2 (⇑, lower row of arrows); parentheses indicate sites that can not be cleaved if branch formation has occurred at nucleotide X. The relative abundances of $X_pC_p^{\ddagger}$ and $C_p^{\ddagger}$ in lanes 5–6, 11–12, 17–18, 23–24, 29–30, 35–36 and 41–42, and of $G_pX_pC_p^{\ddagger}$ and $C_p^{\ddagger}$ in lanes 32–33 and 38–39 reflect the use of the modified position X as the first step nucleophile.

splicing (from Figure 2) at the modified positions (Table II). Second step rates were calculated relative to the amount of first step product available to undergo the second reaction; this resulted in higher rates for the modifications and more accurately reflects the effects on step two. Consistent with the visual appearance in Figure 2, pre-mRNAs with 2,6-diaminopurine and purine branch sites were decreased in splicing only 40 and 20%, respectively, in the first step but underwent the second step at the same rate as the adenine branch site. In contrast, the hypoxanthine branch site was reduced by ~65% for step one, and strongly blocked for step two, which was reduced by >80%; this was less of a step one effect than for the guanine branch site (reduced 90%), but a very similar effect on step two (guanine reduced 70%). Cytosine and uracil branch sites, although similar for step one (reduced ~70%), had differing effects for step two—cytosine was actually better for step two than even adenine, while uracil was decreased ~40%.

These data suggest that for the first step the exocyclic C6-NH_2 group contributes a small positive effect (compare adenine with purine in Table II), whereas the introduction of an amine group onto position 2 has a slight negative effect (compare 2,6-diaminopurine with adenine, and guanine with hypoxanthine) and the C6-oxo/N1-H is inhibitory (compare hypoxanthine with purine). For the second step, it is only the C6-oxo/N1-H of guanine that is inhibitory (compare hypoxanthine with purine, and guanine with 2,6-diaminopurine). Since pre-mRNAs with purine and 2,6-diaminopurine branches underwent step

two efficiently, the C6-NH_2 group does not contribute and an exocyclic C2-NH_2 addition does not interfere, suggesting that the loss of the basic nature of N1 of hypoxanthine and guanine accounts for most of the inhibitory effect on the second step. These results suggest that, although N1 is important for both steps, constituents of the adenine are recognized differently for the two steps of splicing.

Formation of spliceosomes on branch-site base-modified E1–IVS–E2 pre-mRNAs

Recognition of the branch-site base during the first step of splicing could be important during spliceosomal assembly and/or for the chemical reaction. The assembly of splicing complexes on the branch-site modified pre-mRNAs was therefore examined. Samples of the splicing reactions shown in Figure 2 were analyzed on non-denaturing gels (Figure 4), which separate U2 snRNP-containing complex A and spliceosomal complexes B and C from the non-specific H complex. Surprisingly, all of the branch-site modified pre-mRNAs formed the splicing complexes A, B, and C almost as efficiently as the adenine control. Despite a 10-fold decrease for the first chemical step, there was less than a 25% difference in complex formation between an adenine branch site and the guanine branch site (Table II, columns A, B and C; note that the guanine reactions in Figure 4, lanes 33–40, contain fewer counts than the adjacent hypoxanthine and pyrimidine reactions or than the purine reactions). Thus, the dramatic differences in rates of lariat formation (Figure 2) must reflect an interaction required after splicing complex assembly at or near to the time of the first step chemistry.

Table I. Summary of utilization of alternative branch positions

Sequence UGCUGXC	Total yield of lariats	Percentage position X	Percentage other positions[a]
Dap	0.74	94	6.5
A	1.0	97	2.9
Pur	0.99	95	5.3
Hyp	0.45	71	29
G	0.15	41	59
C	0.45	91	9.0
U	0.61	81	19

Utilization of alternative branch positions was determined using substrate pre-mRNAs containing site-specific [^{32}P]phosphates and the sequence and modifications shown. These pre-mRNAs were spliced *in vitro* and the products of the reactions were separated and purified. IVS–E2 and IVS RNAs were digested using RNases and analyzed in polyacrylamide gels (shown in Figure 3). Percentage branch formation at position X was determined as the ratio of branched trinucleotides $X_pC_p^{\ddagger}$ to the total ($X_pC_p^{\ddagger} + C_p^{\ddagger}$), summed for both first step products (IVS–E2 and IVS RNAs, normalized to the relative yield of each lariat species). The amount of RNA in each band was quantitated using a Molecular Dynamics PhosphorImager. Dap, 2,6-diaminopurine; A, adenine; Pur, purine; Hyp, hypoxanthine; G, guanine; C, cytosine; U, uracil. [a]Other positions mostly consisted of the 5′-adjacent guanosine (see Query *et al.*, 1994); this was also evaluated by comparison of $G_pX_pC_p^{\ddagger}$ and $G_pX_pC_p^{\ddagger}$ to $G_pX_pC_p^{\ddagger}$ after digestion using RNase A in Figure 3 (lanes 2–3, 8–9, 14–15, 20–21 and 26–27). Even in the sequence with the highest level of 'cryptic' branches, this does not represent strong activation of other sites, because the overall efficiency of splicing was low (compare total yield of lariats).

Formation of A3′ complexes on branch-site base-modified 3′ partial substrate RNAs

Formation of complexes at the branch site are strongly stimulated on the full-length pre-mRNAs by recognition of the 5′ splice site and were only minimally affected by branch-site modifications. It was possible that formation of complexes on 3′ partial substrate RNAs, lacking a 5′ exon and 5′ splice site, would be a more sensitive assay for recognition of the branch-site adenine. These 3′ partial RNAs form U2 snRNP complexes but do not proceed further in complex formation or undergo any of the chemical steps (Konarska and Sharp, 1986). In fact, this was the case: in contrast to the minor differences observed in complex assembly using full-length E1–IVS–E2 pre-mRNAs, 3′ partial RNAs showed significant differences in A3′ complex formation due to modifications of the branch-site nucleotide (Figure 5 and Table II, column A3′). While 2,6-diaminopurine and purine branch sites were reduced only slightly in formation of the U2 snRNP-containing A3′ complex compared with adenine (~20%), A3′ complexes formed on guanine and hypoxanthine branch sites were reduced by 60–70%, and those on cytidine and uracil sites were decreased by >95 and >80%, respectively.

Table II. Relative rates and yields for splicing at the modified branch-site position

Branch site base X	Relative rates at X			Relative complex formation				
	1st step	2nd step	Normalized 2nd step	Amin	A3′	A	B	C
Dap	0.60	0.63	1.0	0.75	0.88	0.97	0.82	0.76
A	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Pur	0.80	0.83	1.0	0.61	0.83	1.1	1.1	1.1
Hyp	0.34	0.09	0.2	0.04	0.30	1.0	1.0	1.4
G	0.10	0.06	0.3	0.04	0.38	0.86	0.75	0.84
C	0.31	0.44	1.4	0.01	0.03	0.85	0.66	0.59
U	0.32	0.21	0.6	0.03	0.16	0.97	0.79	0.93

Relative rates for the first and second steps of splicing and yields of splicing complexes using pre-mRNAs site-specifically modified at the branch site. First and second step rates were calculated based on the specific nucleotide of branch linkage (see Table I) and normalized to the adenine branch-site reaction; thus they do not directly reflect the appearance of splicing products in Figure 2. Normalized second step is the amount of second step relative to the amount of first step product and thus represents the second step rate independent of first step effects. Relative complex formation was determined as the maximum amount of A, A3′ or Amin complex formed during a time course relative to the input RNA and normalized to the respective adenine-containing RNA. For complex A, the total complexes formed was used (i.e. the sum of A, B and C complexes); for complex B, the total B was used (sum of B and C complexes).

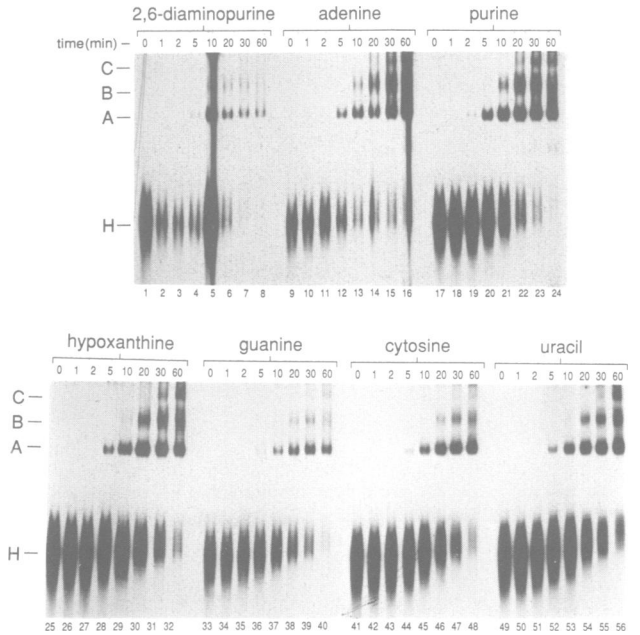

Fig. 4. Time course of splicing complex assembly for branch-site modified E1–IVS–E2 pre-mRNAs. Singly labeled E1–IVS–E2 pre-mRNAs containing the branch sequence 5′-...UGCUGXC...-3′ with 2,6-diaminopurine (lanes 1–8), adenine (lanes 9–16), purine (lanes 17–24), hypoxanthine (lanes 25–32), guanine (lanes 33–40), cytosine (lanes 41–48) or uracil (lanes 49–56) at position X were incubated under splicing conditions for the times indicated (samples from the reactions shown in Figure 2), adjusted to 0.5 mg/ml heparin and loaded onto a native 4% polyacrylamide gel. C, spliceosomal complex C containing U2/5/6 snRNPs and splicing intermediates; B, spliceosomal complex B containing U2/4/5/6 snRNPs and pre-mRNA; A, U2 snRNP complex containing pre-mRNA; H, nonspecific complexes.

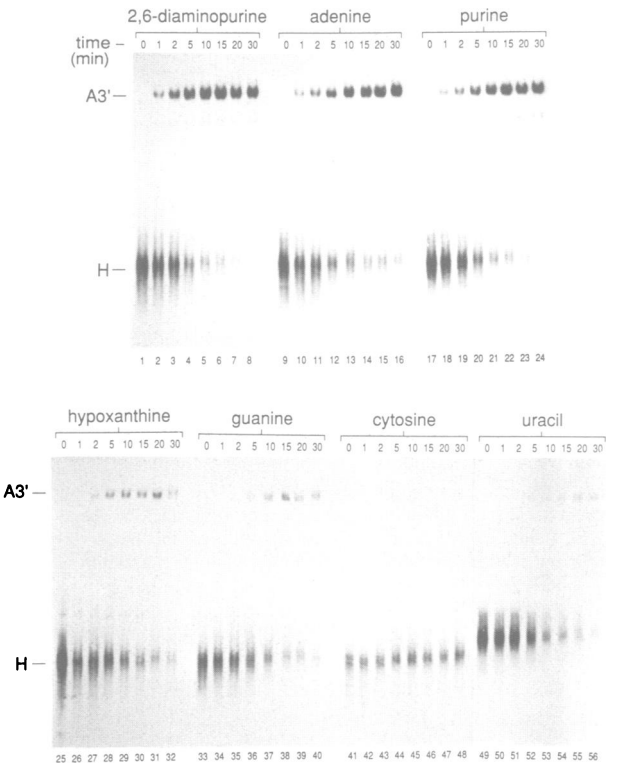

Fig. 5. Time course of complex assembly for branch-site modified 3′ partial substrate RNAs. Singly labeled RNAs containing the branch-site base 2,6-diaminopurine (lanes 1–8), adenine (lanes 9–16), purine (lanes 17–24), hypoxanthine (lanes 25–32), guanine (lanes 33–40), cytosine (lanes 41–48) or uracil (lanes 49–56) were incubated under splicing conditions for the times indicated, adjusted to 0.5 mg/ml heparin and loaded onto a native 4% polyacrylamide gel. A3′, U2 snRNP complex formed on 3′ partial substrate RNAs lacking a 5′ splice site; H, nonspecific complexes.

Formation of minimal A complexes on branch-site base-modified RNA oligonucleotides

To develop an even more sensitive assay for branch-site recognition, we made deletions in the RNA to obtain the shortest segment active in formation of a stable U2 snRNP complex. This RNA oligonucleotide, determined by a systematic deletion study, contained only a branch site and polypyrimidine tract. The minimal A complex formed on this ribo-oligonucleotide, termed Amin complex, contained U2 snRNP and required the branch sequence and the polypyrimidine tract, but was ATP- and U1 snRNP-independent (C.C.Query and P.A.Sharp, in preparation). This Amin complex formed on 2,6-diaminopurine and purine branch-site ribo-oligonucleotides with a 25–40% lower efficiency than on the adenine control (Figure 6 and Table II). Complex formation on hypoxanthine and guanine branch-site ribo-oligonucleotides was decreased by 95%. Amin complex formation on cytosine branch-site ribo-oligonucleotides was nearly undetectable, being decreased by >99%, while that on uracil branch-site ribo-oligonucleotides was decreased by 97%. These results demonstrate that specific recognition of the branch-site adenine occurs early in splicing complex assembly and contributes to the formation and/or stability of the U2 snRNP complex.

Branch-site base-specific cross-links to p14

Previously, three proteins—p14, p35 and p150—were photocross-linked to the branch site as components of complex A as well as A3′, using a linker and photoactive agent that could sample distances up to 15 Å (MacMillan *et al.*, 1994). Similar proteins are found in Amin complexes (C.C.Query and P.A.Sharp, in preparation). It was impractical to combine the previously used 15 Å photochemistry with the base modifications at the branch site. To detect proteins in immediate contact with RNA at the modified branch site, samples were directly irradiated with UV light (Figure 7). RNA oligonucleotides were prepared with the base modifications at the branch position and singly ^{32}P-labeled 3′ to the branch-site position for the 2,6-diaminopurine, adenine, purine and hypoxanthine branch-site RNAs (although attempted, guanine, cytosine and uracil branch-site RNAs do not form enough complex to make this analysis feasible; see Figure 6). These RNAs were incubated to form Amin complex and UV-irradiated; the complexes were separated on a native gel (as in Figure 6) and the Amin complexes were excised. An equivalent amount of each Amin complex was digested with RNase A, and the proteins cross-linked to the labeled RNA fragment were analyzed on SDS–polyacrylamide gels.

With adenine at the branch position, two polypeptides, p14 (a doublet) and a p70, were weakly cross-linked (Figure 7A, lane 6). Neither of these proteins was detected in H complex, where primarily a p60 protein was observed (lane 2). When the base at the branch position was modified, the efficiency of the cross-link to p14 changed

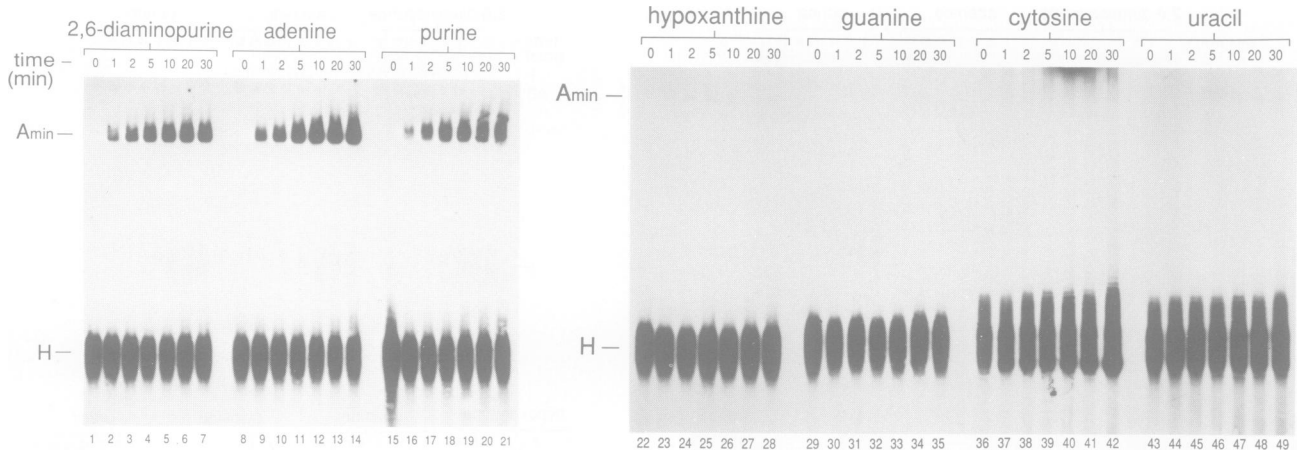

Fig. 6. Time course of complex assembly for RNA oligonucleotides containing only a branch sequence and polypyrimidine tract. RNAs containing the branch-site base 2,6-diaminopurine (lanes 1–7), adenine (lanes 8–14), purine (lanes 15–21), hypoxanthine (lanes 22–28), guanine (lanes 29–35), cytosine (lanes 36–42) or uracil (lanes 43–49) were incubated for the times indicated as described in Materials and methods, adjusted to 0.5 mg/ml heparin and loaded onto a native 4% polyacrylamide gel. Amin, a minimal U2 snRNP complex; H, nonspecific complexes.

greatly (>20-fold) but cross-linking to p70 varied only slightly (~2-fold; compare lanes 5–8). Surprisingly, the cross-link to p14 was more efficient with 2,6-diaminopurine and purine branch-site RNAs. This difference in cross-linking most likely reflects different intrinsic cross-linking efficiencies of the modified bases. The p14 species is almost certainly one of the three proteins mentioned above, previously identified by cross-linking as located within 15 Å of the branch-site nucleotide. The p70 was not detected in the previous study, although a similar molecular weight protein has been detected previously by direct UV cross-linking to the branch-site region (P.Champion-Arnaud and R.Reed, personal communication).

Which base is cross-linked to p14? Digestion of 2,6-diaminopurine and adenine branch-site RNAs with RNase A should generate a trinucleotide containing the labeled phosphate $G_pX_pC_p^{\ddagger}$, as shown in Figure 1B). The influence of the chemical structure of the branch-site base X (especially 2,6-diaminopurine versus adenine) on the efficiency of cross-linking to p14 suggested this base as the probable site of cross-linking. To obtain additional evidence of direct cross-link of p14 to the branch nucleotide, RNA containing 8-azido-adenine at this position was prepared; this RNA was incubated to form Amin complex, irradiated at 302 nm and the cross-linked proteins analyzed as above. Since azide groups absorb UV light at ~302 nm and unmodified adenine absorbs very little at this wavelength, we would expect a direct adenine cross-link to be seen with 8-azido-adenine RNA and not with unmodified adenine RNA. p14 did cross-link to the 8-azido-adenine branch site within Amin complex (Figure 7B, lane 4), whereas the adenine branch-site RNA also irradiated at 302 nm did not generate a labeled p14 product (lane 2). This demonstrates that the site of p14 linkage must be at the 8-azido-adenine, the position of the branch-site nucleotide. In contrast, the p70 species, as well as the p60 species in complex H, showed the same low level of cross-linking with either RNA, suggesting that neither protein is in direct contact with the adenine, but instead might be cross-linked to an adjacent non-adenine position that would be equally activated in either RNA.

The specific contact of the branch site with p14 was also detected in complex A formed on full-length pre-mRNA and was maintained in mature spliceosomes (Figure 7C). E1–IVS–E2 pre-mRNA containing 2,6-diaminopurine at the branch position was prepared as before. This pre-mRNA was incubated under splicing conditions, UV irradiated, and the complexes A, B and C separated (as in Figure 4, lanes 1–12). Cross-linked p14 was again detected in complex A as well as in complexes B and C. These results demonstrate that p14 cross-links in a base-specific fashion and is in intimate contact with the branch-site adenine in the U2 snRNP-containing complex A and also in the spliceosome.

Discussion

The chemical groups on the adenine base at the branch site are recognized during at least three different steps in the splicing of pre-mRNA. These groups were varied systematically and their relative importances in reactions forming early spliceosome complexes and in executing the first and second steps of splicing were determined. Recognition of constituents of the adenine is critical in formation of a U2 snRNP-containing complex on a minimal branch-site oligonucleotide. The efficiencies of the cleavage at the 5′ splice site and lariat formation, and the cleavage and ligation reaction at the 3′ splice site, were differentially affected by changing individual functional groups of the adenine. Moreover, we detected a protein, p14, which interacts with the adenine in a base-specific fashion and may mediate early recognition of this base.

Three reactions dependent on recognition of adenine

Base modifications at the branch site were tested for spliceosome assembly, utilization as the first step nucleophile and continuation through the second step. Specific recognition of adenine contributes to early complex formation: Amin complexes that contain U2 snRNP and form on short RNAs containing the branch sequence and polypyrimidine tract are strongly dependent on the constituents of the base at the branch position. Formation of the U2

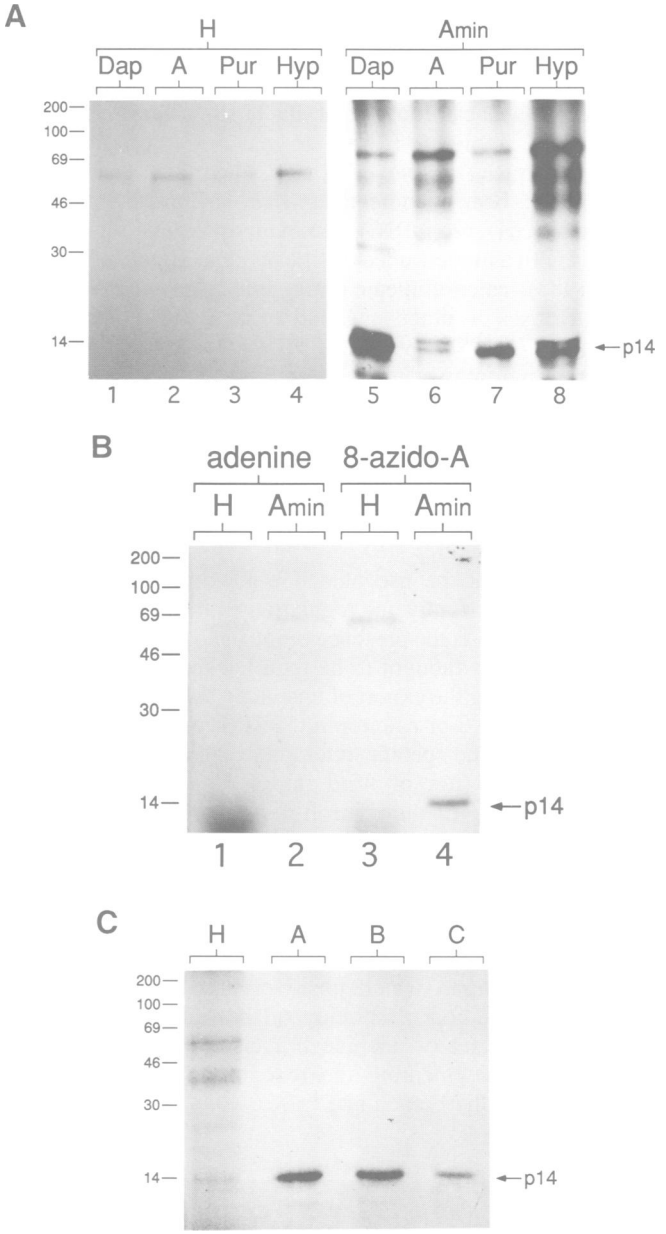

Fig. 7. Two-dimensional analyses of UV cross-linked proteins from splicing complexes. (**A**) Proteins were cross-linked at the branch site using 254-nm light and an RNA oligonucleotide containing only the branch sequence and a polypyrimidine tract and the branch-site base modification indicated above the lanes: 2,6-diaminopurine (Dap, lanes 1 and 5), adenine (A, lanes 2 and 6), purine (Pur, lanes 3 and 7) or hypoxanthine (Hyp, lanes 4 and 8). Amin complexes were separated on native gels (as in Figure 6). Isolated complexes were subsequently treated with RNase and the proteins separated on 16% polyacrylamide gels. (**B**) Amin complexes were formed on adenine branch-site RNA (lanes 1–2) or 8-azido-adenine (8-azido-A) branch-site RNA (lanes 3–4), irradiated at 302 nm and the cross-linked proteins were analyzed as in (A). (**C**) Full-length pre-mRNA containing 2,6-diaminopurine at the branch site was incubated under splicing conditions, irradiated at 254 nm, then separated on native gels (as in Figure 4). Isolated splicing complexes were subsequently treated with RNase and the proteins separated as above. See legend to Figure 4 for abbreviations.

snRNP complex on a complete substrate is not nearly as dependent on these constituents, probably because signals mediated by recognition of the 5′ splice site stabilize U2 snRNP complex formation in spite of branch-site identity. Early complex assembly, as assayed by Amin complexes, is dependent on the exocyclic C6-NH_2 group, which contributes a significant positive effect, while a C2-NH_2 group has a significant negative effect and the C6-oxo/N1-H of guanine is strongly inhibitory. The pyrimidines cytosine and uracil at the branch site are both very inefficient in this assay.

Modifications of the constituents of the branch-site adenine in full-length pre-mRNAs result in significant changes in step one chemistry. However, as mentioned above, the full-length pre-mRNAs form all the splicing complexes efficiently; thus, the effect of branch-site modifications on step one of splicing must occur after spliceosomal assembly. For the first chemical step, the C6-NH_2 group of adenine contributes a small positive effect, and modification or substitution of other chemical substituents have slight to moderate negative effects. In contrast to the Amin reaction where substitutions of either hypoxanthine, cytosine or uracil were strongly inhibitory, 20-fold, these changes only reduced the rate of the first step by 3-fold. Thus, recognition of the adenine for early complex formation can be distinguished from that for the first chemical reaction.

The constituents of the branch-site adenine are recognized differently for the first and second chemical steps of splicing. In contrast to the effects on step one, the presence of an exocyclic C6-NH_2 group does not affect the second step and the C6-oxo/N1-H of guanine is strongly inhibitory. The C6-oxo group should occupy similar space as the C6-NH_2 group and thus probably does not inhibit the second step because of steric considerations; similarly, much bulkier groups placed at position C6 have not inhibited more than 2- or 3-fold [*N*6-hexylamine; C.C.Query and P.A.Sharp, unpublished; *N*6-ethylthiol or *N*6-ethyl-succinamidobenzophenone (MacMillan *et al.*, 1994)]. This focuses attention on the chemical constituents at N1, which is the only adenine-specific position that may contribute positively to the second step. It is possible, then, that the adenine N1 is a specific hydrogen-bond acceptor important for the second step. Although it is difficult to compare directly the groups on the pyrimidines cytosine and uracil with those of the purine analogs, it is notable that cytosine, which undergoes step two efficiently (Table II), has a hydrogen acceptor at position N3, whereas uracil, which does not undergo step two, has a hydrogen donor at N3-H. These chemical groups are roughly comparable with the N1 of adenine and N1-H of guanine, respectively. It is interesting to speculate that the hydrogen-bond acceptor properties of the N1 position of adenine could be involved either in the transesterification chemistry of the second step or in a hydrogen bonding interaction to another base important for the second step active site.

Specific recognition of the adenine base contributes to three different processes during splicing: to early complex formation and to the first and the second steps (summarized in Figure 8). The finding of three adenine recognition events is consistent with the previous results of three differing sets of proteins observed near to the branch nucleotide during spliceosome assembly (MacMillan *et al.*, 1994). The differences in chemical group recognition of the adenine for the first versus the second reactions provides further evidence for the notions that there are significant active site rearrangements in the spliceosome between the first and second chemical events (Moore

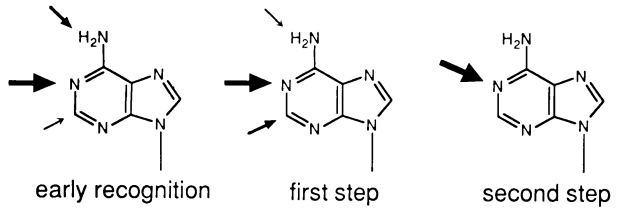

Fig. 8. Schematic summarization of differing recognition of the branch-site adenine. (Left) Early recognition, at or prior to the time of U2 snRNP complex formation, is dependent on N1 as an H-bonding acceptor, the presence of the C6-NH_2 group (removal or substitutions are inhibitory) and substitutions at C2 are slightly inhibitory (see Table II, column 'Amin'). (Center) Step one chemistry is dependent on N1 as an acceptor, substitutions at C2 are inhibitory and the C6-NH_2 group contributes a small positive effect (see Table II, column '1st step'). (Right) Step two chemistry is dependent only on N1 as an acceptor; removal of the C6-NH_2 or substitutions at C2 have no detectable effects specific to this step (see Table II, column 'normalized 2nd step').

and Sharp, 1993) and that the branched structure may participate in forming the second reaction site.

Mediation of early branch-site nucleotide recognition

Using a novel methodology to functionalize site-specifically the pre-mRNA with the photoreagent benzophenone (MacMillan *et al.*, 1994), we previously probed the association of proteins to within 15 Å of the branch-site region during pre-mRNA splicing. Three sets of proteins were distinguished by the kinetics of their associations with the branch site. An early U1 snRNP-dependent cross-link to a p80 species was followed by cross-links to p14, p35 and p150 polypeptides associated with the U2 snRNP–pre-mRNA complex. Concomitant with formation of the spliceosome, a rearrangement of protein factors about the branch region occurred in which the p35 and p150 cross-links were replaced by p220 and p70 species. These results established that the region of the branch site is recognized in a dynamic fashion by multiple distinct proteins during the course of spliceosomal assembly.

Direct UV cross-linking with the branch-site base-modified RNAs used here identified a factor whose interaction was dependent on the identity of the branch-site nucleotide. The existence of such a branch nucleotide recognition factor has been suggested from previous studies in which the branch nucleotide was mutated (e.g. Hornig *et al.*, 1986; Couto *et al.*, 1987; Noble *et al.*, 1988), although early recognition and chemical steps could not be distinguished. This p14 species is in intimate contact with the branch-site adenine, cross-links in a base-specific fashion and thus may determine branch nucleotide recognition in early complex formation. The p14 species is very likely the same as that identified in the previous study as a component of the U2 snRNP–pre-mRNA complex. Since the cross-linking in this case is across a single bond length, we propose that the p14 protein binds directly to the adenine base. The specific interaction of p14 with the branch-site base may account for the sensitivity to adenine modification of the association of U2 snRNP with the minimal sequences of the branch region and polypyrimidine tract.

Natural occurrence of non-adenine branch-site nucleotides

Although the branch-site adenine is highly conserved, other nucleotides can function at this position inefficiently *in vitro* (e.g. Table II; Hornig *et al.*, 1986). In several natural pre-mRNAs, use of a non-adenine branch-site nucleotide has been documented *in vivo*—in human growth hormone (hGH) pre-mRNA removal of the first intron proceeds with a branch at a cytosine residue (Hartmuth and Barta, 1988); calcitonin/calcitonin-gene-related peptide I (CGRP-I) uses a uracil nucleotide in its third intron (Adema *et al.*, 1988, 1990); the HIV-1 Tat/rev transcript second intron primarily utilizes a uracil residue as a branch nucleotide (Dyhr-Mikkelsen and Kjems, 1995) and in SV40 large T antigen pre-mRNA, a cytosine residue is sometimes used as branch position (Noble *et al.*, 1988). In addition, early description of branched structures in nuclear RNA suggested that cytosine may be used *in vivo* (Wallace and Edmonds, 1983).

It is likely that the use of a non-adenine branch site provides an opportunity for regulation in these and other systems. For HIV-1, the presence of intrinsically inefficient splicing signals is thought to be required for the function of Rev in allowing the export of unspliced message (Chang and Sharp, 1989). For calcitonin/CGRP-I, the third intron is subject to tissue-specific alternative splicing to make either calcitonin (use of third intron 3′ splice site) or CGRP-I (skipping of third intron 3′ splice site; van Oers *et al.*, 1994). Although these examples constitute a small set, it is notable that the non-adenine branches are either cytosine or uracil, which are inefficiently active for step one but continue through step two relatively well. These may be desirable characteristics for making an inefficiently spliced intron, upon which positively acting regulators could act. Regulation of splicing of these introns versus the background set of adenine branch introns could be manifested by augmenting the recognition of the non-adenine base at the branch site.

Materials and methods

RNA transcription and synthesis of substrates

pPIP85.B is a modification of pPIP85.A (Moore and Sharp, 1992) to include a single adenosine in the branch region and encodes the following 234 nucleotide sequence: 5′-GGGCGAAUUCGAGCUCACUCUCUU-CCGCAUCGCUGUCUGCGAGGUACCCUACCAG↓GUGAGUAUG-GAUCCCUCUAAAAGCGGGCAUGACUUCUAGAGUAGUCCAG-GGUUUCCGAGGGUUUCCGUCGACGAUGUCAGCUCGUCUCG-AGGGUGCUGACUGGCUUCUUCUCUCUUUUUCCCUCAG↓GUC-CUACACAACAUACUGCAGGACAAACUCUUCGCGGUCUCUGC-AUGCAAGCU-3′. Arrows indicate the 5′ and 3′ splice sites, and A indicates the preferred branch site. RNA(1–145) was transcribed from plasmid pPIP85.B cleaved with *Xho*I. Template for RNA(156–234) was generated by PCR with one deoxyoligonucleotide primer that inserted a T7 promoter immediately 5′ to position 157 and another whose 5′ end corresponded to position 234. These transcriptions were performed under standard conditions (see Query *et al.*, 1994), except that RNA(156–234) was initiated with UpG to insert U(156).

Two- and three-way RNA ligation reactions and gel purification of products were performed as described previously (Moore and Sharp, 1992; Query *et al.*, 1994). Branch site-modified E1–IVS–E2 pre-mRNAs were prepared by joining RNA(1–145), a branch region decamer [RNA(146–155): 5′-GGGUGCUGXC-3′; where X denotes either the base 2,6-diaminopurine, adenine, purine, hypoxanthine, guanine, cytosine or uracil] and 5′-^{32}P-phosphorylated RNA(156–234) using a complementary deoxyoligonucleotide [cDNA(169–136): 5′-GAGAGAAGAA-GCCAGTCAGCACCCTCGAGACGAG-3′]. Synthesis of modified

nucleoside phosphoramidites has been previously described (Strobel *et al.*, 1994). Branch region decamers were prepared by chemical synthesis as described (Strobel *et al.*, 1994) and purified on 20% polyacrylamide gels except for those containing cytosine or uracil at the branch position, which were made by transcription using T7 RNA polymerase as described (Query *et al.*, 1994) and purified on 20% polyacrylamide (19:1), 8 M urea gels run in 1× TBE (89 mM Tris–borate, 2 mM EDTA). 3′ Partial substrate RNAs, lacking a 5′ exon or 5′ splice site, were prepared by joining the branch region decamer (as above) and 5′-[^{32}P]RNA(156–234) using cDNA(169–136). Short ribo-oligonucleotides [RNA(146–179)] containing only a branch sequence and polypyrimidine tract were prepared by joining a branch region decamer and 5′-[^{32}P]RNA](156–179) using cDNA(169–136). Two- and three-way ligation products were purified on 15% polyacrylamide (29:1), 8 M urea gels run in 1× TBE. RNA(146–179), containing 8-azido-adenine or adenine and used for photocross-linking with 302 nm light, was prepared by T7 transcription in the presence of 8-azido-ATP (Sigma) or ATP, respectively, and [α-^{32}P]CTP and purified by Bio-Gel P2 gel-filtration (Bio-Rad) under reduced light conditions.

In vitro splicing assays and analyses of branch-site positions

High specific activity E1–IVS–E2 pre-mRNAs, labeled site-specifically 3′ to the branch site, were incubated under splicing conditions as previously described (Grabowski *et al.*, 1984) and the products were separated and purified on 15% polyacrylamide (29:1) 8 M urea gels run in 1× TBE. The isolated pre-mRNA (E1–IVS–E2) and lariats (IVS–E2 and IVS) were digested for 30 min at 37°C, using 0.06 U/ml RNase T2 (Calbiochem) in 30 mM NaOAc (pH 5.2) and 0.01 mg/ml *Escherichia coli* tRNA; or using 0.5 mg/ml RNase A (BMB) in 10 mM Tris–HCl (pH 7.5), 1 mM EDTA and 0.01 mg/ml *E.coli* tRNA. Products were separated by electrophoresis through a 25% (19:1) polyacrylamide, 8 M urea gel, run in 1× TBE. E1–IVS–E2 pre-mRNAs were analyzed prior to splicing (Figure 1B) using RNases as above or using 0.5 mg/ml RNase A and 1 U/ml RNase T1 (Calbiochem) in 10 mM Tris–HCl (pH 7.5), 1 mM EDTA and 0.01 mg/ml *E.coli* tRNA.

Polyacrylamide gels were quantitated using a Molecular Dynamics PhosphorImager and ImageQuant software version 3.22. For each band in every lane an individual background value was determined from the area in the same lane immediately above or below that band. The relative amount of RNA in each band was expressed as a percentage of the total obtained by summing the values of E1–IVS–E2, IVS–E2 and IVS at that time point. Initial rates were chosen from the linear portion of each curve (four time points); first and second step rates were then calculated based on the specific site of branch formation (see Table I) and normalized to the adenine branch-site reaction. Normalized second step is the amount of second step relative to the amount of first step product.

Native gel analysis of splicing complexes

For E1–IVS–E2 pre-mRNAs, a portion of each splicing reaction described above was adjusted to 0.5 mg/ml heparin and separated by electrophoresis through non-denaturing 4% (80:1) polyacrylamide gels run in 50 mM Tris–glycine. 3′ Partial substrate RNAs were incubated and analyzed using the same conditions and gels. For the analyses of Amin complexes, ATP and creatine phosphate were omitted from the splicing mixes, which were preincubated for 15 min at 30°C to deplete endogenous ATP and adjusted to 10 mM EDTA. Short ribo-oligonucleotides [RNA(146–179), described above] were then added and incubated at 30°C for the times indicated. Reactions were adjusted to 0.5 mg/ml heparin and separated by electrophoresis through non-denaturing 4% (80:1) polyacrylamide gels.

Non-denaturing gels were quantitated as described above and relative complex formation was determined as the maximum amount of A, A3′ or Amin complex formed during a time course relative to the input RNA and normalized to the respective adenine-containing RNA. For complex A, the total complexes formed was used (i.e. the sum of A, B and C complexes); for complex B, the total B was used (sum of B and C complexes).

Photocross-linking assays

High specific activity substrates (10^6 c.p.m./reaction) were incubated in HeLa nuclear extract under splicing conditions as previously described (Grabowski *et al.*, 1984) except that RNasin was omitted (Figure 7C) or under conditions for Amin complex formation (described above; Figure 7A and B); or, for the 8-azido-adenine branch-site RNA, substrates were incubated in HeLa nuclear extract prepared with 0.1 mM 2-mercaptoethanol and no dithiothreitol (DTT) (Figure 7B). Cross-linking was performed on ice by irradiation with a 254 nm wavelength lamp (0.12 W/cm^2 at 1 cm; Ultraviolet Products) for 60 min or, for the 8-azido-adenine branch-site RNA, with a 302 nm wavelength lamp (Ultraviolet Products) for 20 min. Reactions were then digested with 0.02 mg RNase A (Boehringer) or 20 units RNase T1 (Calbiochem) for 30 min at 30°C and analyzed on an 16% (200:1) SDS–polyacrylamide gel.

To analyze proteins from splicing complexes, photocross-linked splicing reactions were prepared as above except that they contained 10^7 c.p.m./reaction and were adjusted to 0.5 mg/ml heparin prior to UV irradiation. Reactions were separated on 4% (80:1) native polyacrylamide gels (Konarska and Sharp, 1986) and frozen; the individual complexes, visualized by autoradiography, were excised. These were incubated with 0.32 mg RNase A/ml of gel in 125 mM Tris–HCl (pH 6.8) at 37°C overnight, then incubated with SDS loading buffer at 37°C for 2 h and 65°C for 5 min and placed directly onto the stacking gel of a disassembled 16% (200:1) SDS–polyacrylamide gel, which was reassembled and run in 0.25 M Tris (pH 8.3), 0.192 M glycine and 0.1% SDS.

Acknowledgements

We thank D.Bartel, B.Blencowe, J.Crispino, A.MacMillan, P.McCaw and M.Moore for their critical review of this manuscript; M.Siafaca for her patience and ever-present assistance; and the members of the Sharp lab for their continued interest and support. C.C.Q. is supported by Leukemia Society of America postdoctoral fellowship (3075-94). This work was supported by United States Public Health Service grant RO1-GM34277 and RO1-AI32486 from the National Institutes of Health to P.A.S. and partially by a Cancer Center Support (core) grant P30-CA14051 from the National Cancer Institute.

References

Abovich,N., Liao,X.C. and Rosbash,M. (1994) The yeast MUD2 protein: an interaction with PRP11 defines a bridge between commitment complexes and U2 snRNP addition. *Genes Dev.*, **8**, 843–854.

Adema,G.J., Bovenberg,R.A.L., Jansz,H.S. and Baas,P.C. (1988) Unusual branch point selection involved in splicing of the alternatively processed calcitonin/CGRP-1 pre-mRNA. *Nucleic Acids Res.*, **16**, 9513–9526.

Adema,G.J., van Hulst,K.L. and Baas,P.D. (1990) Uridine branch acceptor is a *cis*-acting element involved in regulation of the alternative processing of calcitonin/CGRP-1 pre-mRNA. *Nucleic Acids Res.*, **18**, 5365–5373.

Behrens,S.-E., Galisson,F., Legrain,P. and Lührmann,R. (1993) Evidence that the 60-kDa protein of 17S U2 small nuclear ribonucleoprotein is immunologically and functionally related to the yeast PRP9 splicing factor and is required for the efficient formation of prespliceosomes. *Proc. Natl Acad. Sci. USA*, **90**, 8229–8233.

Blackburn,P. and Moore,S. (1982) Pancreatic ribonuclease. In Boyer,P.D. (ed.), *The Enzymes*. Academic Press Inc., New York, pp. 317–433.

Brosi,R., Gröning,K., Behrens,S.-E., Lührmann,R. and Krämer,A. (1993a) Interaction of mammalian splicing factor SF3a with U2 snRNP and relation of its 60-kD subunit to yeast PRP9. *Science*, **262**, 102–105.

Brosi,R., Hauri,H.-P. and Krämer,A. (1993b) Separation of splicing factor SF3 into two components and purification of SF3a activity. *J. Biol. Chem.*, **268**, 17640–17646.

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

Champion-Arnaud,P. and Reed,R. (1994) The prespliceosome components SAP 49 and SAP 145 interact in a complex implicated in tethering U2 snRNP to the branch site. *Genes Dev.*, **8**, 1974–1983.

Chang,D.D. and Sharp,P.A. (1989) Regulation by HIV Rev depends upon recognition of splice sites. *Cell*, **59**, 789–795.

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

Dyhr-Mikkelsen,H. and Kjems,J. (1995) Inefficient spliceosome assembly and abnormal branch site selection in splicing of an HIV-1 transcript in vitro. *J. Biol. Chem.*, **270**, 24060–24066.

Freyer,G.A., Arenas,J., Perkins,K.K., Furneaux,H.M., Pick,L., Young,B., Roberts,R.J. and Hurwitz,J. (1987) *In vitro* formation of a lariat structure containing a G2′–5′G linkage. *J. Biol. Chem.*, **262**, 4267–4273.

Gaur,R.K., Valcárcel,J. and Green,M.R. (1995) Sequential recognition

of the pre-mRNA branch point by U2AF^{65} and a novel spliceosome-associated 28-kDa protein. *RNA*, **1**, 407–417.
Grabowski,P.J., Padgett,R.A. and Sharp,P.A. (1984) Messenger RNA splicing in vitro: An excised intervening sequence and a potential intermediate. *Cell*, **37**, 415–427.
Hartmuth,K. and Barta,A. (1988) Unusual branch point selection in processing of human growth hormone pre-mRNA. *Mol. Cell. Biol.*, **8**, 2011–2020.
Hornig,H., Aebi,M. and Weissmann,C. (1986) Effect of mutations at the lariat branch acceptor site on β-globin pre-mRNA splicing in vitro. *Nature*, **324**, 589–591.
Jamison,S.F. and Garcia-Blanco,M.A. (1992) An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. *Proc. Natl Acad. Sci. USA*, **89**, 5482–5486.
Konarska,M.M. and Sharp,P.A. (1986) Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell*, **46**, 845–855.
Konarska,M.M. and Sharp,P.A. (1987) Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell*, **49**, 763–774.
Krämer,A. (1995) The biochemistry of pre-mRNA splicing. In Lamond,A.I. (ed.), *Pre-mRNA Processing*. R.G.Landes Company, Austin, TX, pp. 35–64.
Lamm,G.M. and Lamond,A.I. (1993) Non-snRNP protein splicing factors. *Biochim. Biophys. Acta*, **1173**, 247–265.
MacMillan,A.M., Query,C.C., Allerson,C.R., Chen,S., Verdine,G.L. and Sharp,P.A. (1994) Dynamic association of proteins with the pre-mRNA branch region. *Genes Dev.*, **8**, 3008–3020.
Madhani,H.D. and Guthrie,C. (1994) Dynamic RNA–RNA interactions in the spliceosome. *Annu. Rev. Genet.*, **28**, 1–26.
Michaud,S. and Reed,R. (1991) An ATP-independent complex commits pre-mRNA to the mammalian spliceosome assembly pathway. *Genes Dev.*, **5**, 2534–2546.
Moore,M.J. and Sharp,P.A. (1992) Site-specific modification of pre-mRNA: the 2′ hydroxyl groups at the splice sites. *Science*, **256**, 992–997.
Moore,M.J. and Sharp,P.A. (1993) The stereochemistry of pre-mRNA splicing: evidence for two active sites in the spliceosome. *Nature*, **365**, 364–368.
Moore,M.J., Query,C.C. and Sharp,P.A. (1993) Splicing of precursors to mRNA by the spliceosome. In Gesteland,R. and Atkins,J. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 303–357.
Nilsen,T.W. (1994) RNA–RNA interactions in the spliceosome: unraveling the ties that bind. *Cell*, 78, 1–4.
Noble,J.C.S., Prives,C. and Manley,J.L. (1988) Alternative splicing of SV40 pre-mRNA is determined by branch site selection. *Genes Dev.*, **2**, 1460–1475.
Parker,R. and Siliciano,P. (1992) Evidence for an essential non-Watson–Crick interaction between the first and last nucleotides of a nuclear pre-mRNA intron. *Nature*, **361**, 660–662.
Parker,R., Siliciano,P.G. and Guthrie,C. (1987) Recognition of the TACTAAC box during mRNA splicing in yeast involves base pairing to the U2-like snRNA. *Cell*, **49**, 229–239.
Query,C.C., Moore,M.J. and Sharp,P.A. (1994) Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev.*, **8**, 587–597.
Reed,R. and Maniatis,T. (1988) The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes Dev.*, **2**, 1268–1276.
Ruby,S.W. and Abelson,J. (1988) An early hierarchic role of U1 small nuclear ribonucleoprotein in spliceosome assembly. *Science*, **242**, 1028–1035.
Ruskin,B., Zamore,P.D. and Green,M.R. (1988) A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell*, **52**, 207–219.
Rymond,B.C. and Rosbash,M. (1992) Yeast pre-mRNA splicing. In Jones,E.W., Pringle,J.R. and Broach,J.R. (eds), *The Molecular and Cellular Biology of the Yeast Saccharomyces: Gene Expression*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 143–192.
Scadden,A.D.J. and Smith,C.W.J. (1995) Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. *EMBO J.*, **14**, 3236–3246.
Séraphin,B. and Rosbash,M. (1989) Identification of functional U1 snRNP–pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell*, **59**, 349–358.
Séraphin,B. and Rosbash,M. (1991) The yeast branchpoint sequence is not required for the formation of a stable U1 snRNP pre-mRNA complex and is recognized in the absence of U2 snRNA. *EMBO J.*, **10**, 1209–1216.
Staknis,D. and Reed,R. (1994) Direct interactions between pre-mRNA and six U2 small nuclear ribonucleoproteins during spliceosome assembly. *Mol. Cell. Biol.*, **14**, 2994–3005.
Strobel,S.A., Cech,T.R., Usman,N. and Beigelman,L. (1994) The 2,6-diaminopurine riboside•5-methylisocytidine wobble base pair: an isoenergetic substitution for the study of G•U pairs in RNA. *Biochemistry*, **33**, 13824–13835.
Uchida,T. and Egami,F. (1971) Microbial ribonucleases with special reference to RNases T1, T2, N1, and U2. In Boyer,P.D. (ed.), *The Enzymes*. Academic Press Inc., New York, pp. 205–250.
van Oers,C.C.M., Adema,G.J., Zandberg,H., Moen,T.C. and Baas,P.D. (1994) Two different sequence elements within exon 4 are necessary for calcitonin-specific splicing of the human calcitonin/calcitonin gene-related peptide I pre-mRNA. *Mol. Cell. Biol.*, **14**, 951–960.
Wallace,J.C. and Edmonds,M. (1983) Polyadenylated nuclear RNA contains branches. *Proc. Natl Acad. Sci. USA*, **80**, 950–954.
Wu,J. and Manley,J. (1989) Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes Dev.*, **3**, 1553–1561.
Zamore,P.D., Patton,J.G. and Green,M.R. (1992) Cloning and domain structure of the mammalian splicing factor U2AF. *Nature*, **355**, 609–614.
Zhuang,Y. and Weiner,A.M. (1989) A compensatory base change in human U2 snRNA can suppress a branch site mutation. *Genes Dev.*, **3**, 1545–1552.
Zhuang,Y., Goldstein,A.M. and Weiner,A.M. (1989) UACUAAC is the preferred branch site for mammalian mRNA splicing. *Proc. Natl Acad. Sci. USA*, **86**, 2752–2756.

Received on September 21, 1995; revised on October 26, 1995