Interplay between U2 snRNP and 3' splice factor(s) for branch point selection on human β -globin pre-mRNA

Christine Alibert, Jamal Tazi, Jamal Temsamani⁺, Philippe Jeanteur*, Claude Brunel and Guy Cathala

URA CNRS 1191, Laboratoire de Biochimie, CRLC Val d'Aurelle Paul Lamarque 34094 Montpellier cedex 2 et Laboratoire de Biologie Moléculaire, Université des Sciences et Techniques du Languedoc, 34095 Montpellier cedex 2, France

Received October 25, 1989; Revised and Accepted December 18, 1989

ABSTRACT

We investigated the interaction of U2 snRNP with the branch-3' splice site region of three human β -globin pre-mRNAs carrying nearly complete (BamHI RNA), 24 nt (Avall RNA) and 14 nt (Accl RNA) of exon 2. All supported splicing, but mRNAs yields were respectively 2 and 10 times lower for Avall and Accl RNAs than for BamHI. Analysis of RNase T₁-resistant fragments immunoprecipitated by an anti-(U2)RNP antibody at early times of the splicing reaction showed that the protection encompasses both the branch point region and the end of the intron in BamHI and Avall, but essentially only the branch point in Accl RNAs. Later on, this protection becomes less detectable in BamHI, is reinforced in Avall and remains poorly detectable in Accl RNAs. Similar experiments performed at late times with an anti-Sm antibody recognizing all snRNPs showed that the end of the intron is protected in all but BamHI RNAs. These results support the conclusion that U2 snRNP binds to a fully efficient precursor (BamHI RNA) through another factor(s) recognizing the 3' splice site (U5 snRNP and the so-called U2AF protein are likely candidates). Either the absence of an initial contact between U2 snRNP and the factor(s) recognizing the end of the intron (Accl RNA) or the inability of this ternary complex to undergo a conformational change (Avall RNA) could render these severely truncated precursors poor substrates. These different situations have consequences on the branch point selection itself. BamHI and Avall RNAs use three functional branch points at early times, the usual A residue at -37 and two U residues at -17 and -22. Accl RNA uses only one branch point at -37. Later on, all three branch points are used at the same rate in Avall, while the usual one prevails in BamHI RNAs.

INTRODUCTION

A general outline of the process by which pre-mRNAs are spliced has been conclusively established. As a first step during which ATP and Mg⁺⁺ are required, pre-mRNA is assembled into a multicomponent complex called spliceosome (1-6). Once the spliceosome is assembled, splicing proceeds by two separate transesterification reactions.

The abundant U1, U2, U5 and U4-U6 snRNPs (reviewed in ref.7) all appear to be involved in splicing (reviewed in ref. 8-11). Early experiments first established the inhibition of splicing by anti-Sm, anti-RNP and anti-(U2) RNP (12-14) and then by oligonucleotide-directed RNase H cleavage of U1, U2 and U4-U6 snRNPs (15-19). As the highly compact organization of U5 snRNP in nuclear extracts (20, 21) made it unamenable to the latter approach, its involvement could only be inferred indirectly (20). Recent analyses of purified splicing complexes (2, 22-24) support this conclusion. Nuclease protection experiments have also been used to map the binding sites for snRNPs. U1 (16, 25) and U2 (2, 16, 23, 26, 27) bind to the 5' splice site and to the branch point sequence, respectively. Compelling evidence has been obtained from both mammalian and yeast systems that U1 interaction is mediated at least in part by base pairing (28, 29). In *Sacharomyces cerevisiae*, a region of U2 snRNA base pairs with the conserved so-called TACTAAC box surrounding the branch site (30). Although no such stringent requirement seems to exist in higher eukaryotes where cryptic branch sites are used in intron mutants (31, 32), recent results have demonstrated that U2 snRNA also base pairs with a sequence resembling the yeast TACTAAC box in the arrangement of purines and pyrimidines (33-36). U5 snRNP is thought to bind to the 3' splice site (20) via an associated snRNP protein (37, 38). As to U4-U6 snRNP, no specific binding site has been identified although an ATP-dependent association with U5 snRNP may exist (24, 39-41).

In spite of these remarkable findings, two major problems are still a matter of discussion. One concerns the early events of the

* To whom correspondence should be addressed

⁺ Present address: Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545, USA

spliceosome assembly pathway that may precede binding of U2 snRNP. A number of candidates have been proposed. These include U1 snRNP (2, 28, 42–44), hnRNP proteins (45, 46), a U2 snRNP auxiliary factor U2AF (41, 47) and IBP-U5 snRNP (47, 48). Another important problem concerns the snRNP-snRNP interactions that probably trigger splicing decisions within the spliceosome. In this respect, it seems that contacts between U1 and U2 snRNPs may be crucial to position the 5' splice site close to the branch site (49).

In this work, we investigated the interaction of U2 snRNP with a normal human β -globin pre-mRNA as well as two less efficient precursors, derived from the same plasmid, truncated in exon 2 near the 3' splice site. Using RNase T₁ protection-immunoprecipitation assays, we obtained patterns of protected sequences suggesting noticeable differences about the mechanism of U2 snRNP binding to the three precursors. Moreover, branch point analyses performed during the course of splicing of these same RNAs revealed that functional branches are formed on U residues in addition to the authentic one on the -37 A position. These findings are interpreted in terms of an interplay between U2 snRNP and the 3' splice site factor(s).

MATERIALS AND METHODS

Materials

SP6 RNA polymerase, RNasin, restriction enzymes, DNase I, reverse transcriptase, deoxynucleotides and dideoxynucleotides triphosphates were from Promega Biotec, RNase T₁ from Calbiochem, RNase H, intestinal bovine alkaline phosphatase and ribonucleotides triphosphates from Boehringer Mannheim and m⁷G5' pppG5' cap from Pharmacia. α -³²P triphosphates, γ -³²P ATP and α -³²P dATP (specific activity 410 Ci/mmol) were from Amersham. Oligonucleotide primers were synthesized in our laboratory using a Biosearch SAM I apparatus that uses phosphotriester chemistry. All other chemicals were of analytical grade.

Plasmids, SP6 transcription and splicing in vitro

pSP64H β Δ 6 and pSP64H β Δ IVS1,2 (50) were generous gifts from Dr. T. Maniatis (Harvard University, Boston).

High specific activity SP6 RNA transcripts were obtained as previously described (37) and were used as substrates in standard splicing reactions. Quantitation of splicing efficiency was based on Cerenkov counting of bands eluted from the gel. For primer extension analyses, low specific activity precursor RNAs (5–8 $\times$ 10³ cpm/mmol) were synthesized in 50 μ l reaction mixtures containing 4 μ g linearized DNA template, 40 mM Tris-HCl pH 7.5, 6 mM MgCl₂, 10 mM DTT, 2.7 mM spermidine, 40 Units RNasin, 0.5 mM of each unlabeled NTPs, 1 μ Ci α -³²P UTP and 2.5 mM m⁷G5' pppG5'. Upon transcription, the RNAs were purified by gel electrophoresis. They were used in a final volume of 2.5 ml containing 12.8 mM triethanolamine pH 7.9, 3.2 mM MgCl₂, 0.5 mM ATP, 20 mM creatine phosphate, 2.7% polyvinyl alcohol, 14% glycerol, 62 mM KCl, 0.12 mM EDTA, 0.7 mM DTT and 15 μ g of transcript. Intermediate and product RNAs were first separated on a 4.5% polyacrylamide gel and then the bands of interest were purified through another 10% gel. Following autoradiography, individual bands were cut out of the gel and eluted overnight in 300 μ l of 0.7 M ammonium acetate containing 1 mM EDTA and 0.1% SDS. Acrylamide debris were removed by filtration through a minisart-NML filter which was washed with 200 μ l of 0.7 M ammonium acetate. RNA

was precipitated twice in the presence of 2 μ g glycogen as carrier and washed with 70% ethanol.

Immunoprecipitation/RNase T₁ digestion analyses

The patient anti-(U2) RNP (Ya) and the mouse monoclonal anti-Sm (7.13) antibodies were generous gifts respectively from Dr. J. Steitz (Yale University, New Haven) and Dr. Hoch (Agouron Institute, La Jolla). Immunoprecipitation/RNase T₁ digestions from splicing reactions (containing 1 $\times$ 10⁶ cpm Cerenkov of RNA substrate) were as described (16, 20, 49) using the above antibodies. All fractionations of RNase T₁-resistant fragments were on prerun 20% polyacrylamide-8 M urea gels in TBE buffer.

Mapping of branch points

For mapping RNA branch points in purified lariat RNAs, either the intron-exon 2 or the excised intron were incubated with 50–100 pmoles of one of the synthetic DNA oligonucleotides used as primers in 10 μ l of hybridization buffer containing 34 mM Tris-HCl (pH 8.3 at 42°C), 100 mM NaCl, 6 mM MgCl₂ and 5 mM DTT. The reaction mixture was heated to 65°C for 3 min then incubated at 40°C for 90 min before drying in a Speed Vac centrifuge. Reverse transcription was carried out in the same test tube in a final volume of 10 μ l containing in the same buffer as above 1 mM each dCTP, dGTP and dTTP, 20–30 μ Ci α -³²P dATP, 30 units RNasin and 15 units reverse transcriptase. After incubation for 30 min at 41°C, 1 μ l of a mixture containing all cold dNTPs (2 mM each) was added and incubation continued for an additional 30 min. The reaction was stopped by adding 10 μ l of 98% formamide containing 2 mM EDTA then heated at 70°C for 5 min and loaded on a 15% polyacrylamide sequencing gel.

In cases where the sequence had to be determined, the same procedure was applied except that dideoxynucleotides were added (dNTP/ddNTP=2.5/1). The reaction was carried out in 5 μ l instead of 10 μ l and the incubation time was reduced to 15 min.

To confirm that blocks to reverse transcription are due to branched nucleotides, controls were carried out on lariat RNAs previously debranched with the cytoplasmic debranching activity (26). Briefly, purified lariat RNAs were incubated under standard splicing conditions except that HeLa nuclear extract was replaced by an S 100 cytoplasmic fraction prepared according to Dignam et al. (51).

In one experiment we used a deoxyoligonucleotide labeled at its 5' end with γ -³²P ATP as described (52).

Direct RNA analyses

Intron lariat-exon 2 intermediates and introns lariat were recovered from polyacrylamide gels. They were digested by 54 U of RNase T₁ in 20 μ l of 10 mM Tris-HCl (pH 7.4), 1 mM EDTA and 10 μ g of tRNA carrier for 30 min at 37°C. RNase T₁ products were fractionated on prerun polyacrylamide gels, recovered and concentrated with n-butanol and dried in a Speed Vac apparatus. Alkaline hydrolyses and digestions by RNase A, RNase P1 and alkaline phosphatase were as described (53–55). Hydrolysis products were fractionated by two-dimensional thin-layer chromatography on cellulose plates using isobutyric acid-concentrated-NH₄OH-H₂O (66/1/33, v/v/v) in the first dimension and 0.1 M sodium phosphate, pH 6.8-ammonium sulfate-n-propanol (100/60/2, v/w/v) in the second dimension (56).

Figure 1. Characterization of in vitro splicing products and intermediates from the AvaII and AccI human β -globin mRNAs precursors. Panel A: Time course of in vitro splicing. For each precursor, lanes 1 to 4 show the electrophoretic pattern of intermediates and products resolved in a 4.5% polyacrylamide-urea gel at times 0, 20, 40 and 60 min respectively. Lanes 5 show E1-E2 transcripts that are used here as mRNA markers. They were directly generated from pSP64H β Δ IVS1.2 plasmid linearized either with AccI or AvaII restriction enzymes. Intermediates and products were identified by reference to earlier works (BamHI) or are designated using conventional abbreviations. Panel B: Bands derived from AccI and AvaII RNAs were designated according to their decreasing size (a1 to e1 and a2 to e2 respectively), eluted from the gel in the presence of SDS and electrophoresed through a 10% polyacrylamide-urea gel in which marker mRNA (m) is derived from the pSP64H β Δ IVS1.2 plasmid.

RESULTS

Splicing efficiency and accuracy of human β -globin pre-mRNA truncated in the second exon.

Three SP6 RNA precursors labeled on U residues were generated from the pSP64H β Δ 6 plasmid (50) terminating in the second exon 209 (BamHI RNA), 24 (AvaII RNA) and 14 (AccI RNA) nucleotides downstream from the 3' splice site and processed in vitro under standard conditions. The presence of intermediates and products at 0, 20, 40 and 60 min was monitored by electrophoretic analyses on both 4.5% and 10% polyacrylamide gels containing 8M urea. While all intermediates and products from Bam HI RNA could be unambiguously identified in a 4.5% gel (not shown) by reference to previously published works, the presence and the identity of those from the other precursors (Figure 1A) had to be verified by analysing the bands eluted from the first 4.5% gel in a 10% gel (Figure 1B) in which lariat RNAs exhibit a reduced mobility in comparison to linear molecules of the same length. As an example, band c_2d_2 from AvaII RNA (Figure 1A), once eluted from the 4.5% gel could be separated into two species corresponding to exon 1 (158 nucleotides) and intron lariat-exon 2 (154 nucleotides) (Figure 1B).

Regardless of the length of exon 2, cleavage at 5' splice site occurred efficiently, as judged by the presence of quite similar amounts of exon 1 and intron lariat-exon 2 intermediates (respectively bands c_1 - c_2 and d_1 - d_2 in Figure 1A). In this respect, our result agrees perfectly with that of Furdon and Kole (57) in similar studies with identical precursors. While these authors concluded that the ligation of the exons was completely inefficient

in both AvaII and AccI RNAs, we found in contrast that this second step takes place accurately but with approximately a two and ten-fold decrease in efficiency, for AvaII and AccI RNAs respectively, in comparison with BamHI RNA. A similar observation has been also reported (58). Some faint bands of the size of debranched intron-exon 2 molecules appeared during splicing of AccI RNA. They comigrated with band d_1 on the 4.5% gel (Figure 1A) and appeared when band d_1 was electrophoresed on a 10% gel (Figure 1B, lane d_1).

The extent of protection surrounding the branch-3' splice site is different in BamHI, AvaII and AccI RNAs

We investigated the possibility that the relative inefficiency of precursors truncated at a short distance downstream from the 3' splice site may be due to an altered interaction of the snRNPs involved in the recognition of the branch point sequence and the 3' splice site. This was achieved using the RNase T₁ protection-immunoprecipitation assay previously described (20, 49) which allows to determine the sites of interaction between the snRNPs and the pre-mRNA. Our interest in the early events of spliceosome assembly dictated the use of crude splicing extracts rather than of completed spliceosomes isolated on glycerol gradients, even if they contain various types of RNP complexes (pre-spliceosomes, spliceosomes and post-spliceosomes). In any case, most of the RNase T₁ fragments examined here were identical to those previously characterized in isolated spliceosomes (49).

Each ³²P-labeled pre-mRNA was incubated at 30°C in a

Figure 2. Anti-(U2) RNP immunoprecipitation of protected fragments in full-length and truncated human β -globin precursors. BamHI, AvaII and AccI transcripts labeled with $\alpha^{32}\text{P}$ UTP were incubated under standard splicing conditions. At 20, 40, and 60 min, aliquots were withdrawn and immunoprecipitated in the presence of T₁ RNase as described (49). Lanes 1–3 in A, lanes 4–6 in A and lanes 1–3 in C refer to 20–60 min immunoprecipitations from AvaII, BamHI and AccI RNAs, respectively. Lane 4 in C refers to anti-(U2) immunoprecipitated fragments from BamHI RNA following 150 min of incubation. M designates RNA markers obtained either by RNase T₁ hydrolysis of BamHI pre-mRNA or from a mixture of DNA templates transcribed in vitro. In some cases all these markers were mixed. Their sizes are indicated. Panels B and D show RNase T₁ secondary analyses of the main protected fragments isolated in the experiments shown in panels A and C, respectively. In B, fragments were electrophoresed either digested (on the left of the lane showing markers) or undigested (on the right). In panel D, fragments D (lane 1, AccI RNA; lane 3, BamHI RNA) and E (lane 2, AccI RNA; lane 4, BamHI RNA) were digested and electrophoresed.

Figure 3. (A) Position and length of the resistant RNase T₁ fragments along the human β -globin pre-mRNA sequence. The pre-mRNA sequence is derived from the DNA sequence reported (59). The protected regions corresponding to T₁-resistant fragments were designated according to the nomenclature used (49). (B) Schematic representation of the U2 snRNP-protected regions seen in Figure 2. The thickness of each protected fragment represents an estimate of its amount calculated by counting the radioactivity of bands seen in Figure 2 A, C corrected for their uridine content.

splicing extract and RNase T₁ protected fragments were immunoprecipitated by anti-(U2)RNP antibody after 20, 40 and 60 min. After deproteinization, they were separated through a 20% denaturing gel (Figure 2A, C) and the major ones subsequently identified by RNase T₁ (Figure 2B, D). Fragments designated A, B, B'', D, E and F are those previously characterized in similar experiments (20, 49). In addition to the above, new fragments are generated (B1, D1, D2 and F', Figure 2A). B1 may be the 3' dephosphorylated version of B (see Figure 8 below). D1, D2 and F' are shortened versions of D/E and F, respectively. Figure 3 shows the mapping of these fragments on the β -globin pre-mRNA sequence. The situation is most simple with the AccI substrate (Figure 2C) where the previously recognized D and E fragments (49) are most prominent at all times. Quite on the opposite, the patterns of protected fragments with BamHI and AvaII are strongly and differentially affected by the time of reaction. A similar pattern characterized by the absence of bands D and E (Figure 2A, lanes 1 and 4) is observed for both substrates at 20 min. At later times, the BamHI and AvaII patterns strikingly depart from one another with respect to the family of bands (F', B'' and B + B1) which share the 3' splice site 19 mer: they all decrease in BamHI RNA while they accumulate in AvaII RNA (Figure 2A and Figure 3 for a schematic representation). After a prolonged incubation (150 min), bands D and E eventually become prominent in BamHI RNA which now exhibits the same pattern (Figure 2C, lane 4) as observed with AccI RNA from the very beginning.

At least two scenarios could account for these different observations. First, U2 snRNP could interact independently of all other snRNPs and have contacts with both the branch point sequence (fragments D1, D2) and part or all of the polypyrimidine tract (fragment B), except in AccI RNA in which only a weak protection encompassing the polypyrimidine tract is seen with the anti-(U2) RNP antibody (band F). This could be the case in BamHI RNA at 20 min and in AvaII RNA at any time. Second, both U2 snRNP and another factor recognizing the 3' splice site could interact together and be responsible for the protection surrounding the branch-3' splice site.

We performed additional experiments using the anti-Sm monoclonal antibody 7.13 (Figure 4A). Each precursor was

Figure 4. Anti-Sm immunoprecipitations of protected fragments in full-length and truncated human β -globin precursors. All RNA transcripts were labeled with α^{32} P UTP and RNA markers were as in Figure 2. In panel A, immunoprecipitations were following 90 min of incubation. Lanes 1–3 refer to AccI, AvaII and BamHI RNAs, respectively. Panel B shows subsequent RNase T₁ analyses of the protected B fragments seen in panel A. Lanes 1–3 refer to fragment B from AccI, AvaII and BamHI RNAs respectively.

incubated in the splicing extract during 90 min, i.e. under conditions that do not allow immunoprecipitation of fragment B by the anti-(U2) RNP antibody neither in BamHI nor in AccI RNAs. At first glance, the patterns of bands obtained with the three precursors are quite similar and all major bands (H, G, F, F', E, D, D2, B and A) previously characterized (49) have

Figure 5. Localization of stops to reverse transcription in both introns lariat-exon 2 and excised introns generated from the three human β -globin mRNAs precursors. In the top is a schematic diagram positioning the complementary deoxyoligonucleotides used in primer extension analyses of BamHI and AvaII intron lariat-exon 2 splicing intermediates (oligo a) and both AccI intron lariat-exon 2 and AvaII precursor (oligo c). As described in the text, oligo b was used with BamHI intron lariat-exon 2 to generate, by site specific cleavage with RNase H, an RNA having the same size as AvaII intron lariat-exon 2. Positions -1 and $+1$ correspond to the last and to the first nucleotide respectively of the intron and the second exon. Star at -37 position indicates the usual branch point. Sequencing gels were obtained from the following templates: a, AccI intron lariat-exon 2; b, AvaII precursor; c, AvaII intron lariat-exon 2; d, BamHI intron lariat-exon 2; e, AvaII excised intron. Portions of the human β -globin pre-mRNA sequence (59) are shown in which stops to reverse transcriptase are indicated by thick arrows. Other putative stops are indicated by stars.

been identified by fingerprint analyses (not shown). Interestingly, analyses of the 19 nt long fragments (Figure 4B) clearly show the presence of fragment B in all but BamHI RNAs (compare lane 3 with lanes 1 and 2). Another important feature of the anti-Sm antibody assays is the presence of fragment B from AccI RNA, in contrast to the experiment carried out with the anti-(U2) RNP antibody showing that U2 snRNP directly binds to the branch point region. Such a result has two implications in the case of AccI RNA. First, the factor protecting the 3' splice site is a snRNP different from U2 or/and a protein recognized by the anti-Sm antibody. Second, no interaction occurs between the branch point region-bound U2 snRNP and the factor(s) bound

to the end of the intron. Assuming that this (these) same factor(s) different from U2 snRNP also interact(s) with BamHI and AvaII precursors, it seems likely that the anti-U2 RNP immunoprecipitation of fragments containing the 19 mer oligonucleotide is due to an interaction between U2 snRNP and this (these) factor(s) bound to the 3' splice site, rather than the U2 snRNP alone having contacts with both the branch point sequence and the end of the intron.

Presence of multiple branchpoints in AvaII RNA

If a normally efficient pre-mRNA (BamHI RNA) undergoes a change of the protection surrounding the branch-3' splice site,

Figure 6. All stops to reverse transcription occurring in lariat RNAs are authentic branch points. Panel A: Primer extension analysis of both intron lariat-exon 2 (lane 2) and excised intron (lane 3) from *Ava*II RNA using 5' end labeled primer c. Lane M shows DNA markers. Panel B: Intron lariat-exon 2 from *Ava*II RNA was debranched as described in Materials and Methods prior to primer extension using internal labelling in the absence of dideoxynucleotides. The cDNA product was electrophoresed in a 15% polyacrylamide-urea gel (lane 1) and detected by autoradiography in comparison to DNA markers from pGEM 1 DNA cleaved with *MS*pI (lane M) and to cDNAs products obtained as in Figure 3 from the corresponding unbranched lariat RNA (lanes 2). Numbers on the left indicate the size of the cDNAs products. Panel C: Schematic drawing of the RNA template, the 16 nt primers a and c and the cDNA products (length indicated on the right) and the positions of branch points (numbers above the sequence).

then the different situations seen in our protection experiments with the two other precursors probably indicate an altered organization of the corresponding spliceosomes. A possibility would be that unusual branch points may be generated leading to intron lariat-exon 2 RNAs not normally used in the second step of the reaction or used with a poor efficiency. We therefore purified intron lariat-exon 2 species produced from the three RNAs at 60 min and mapped them for the possible presence of unusual branch points in addition to the normal one.

a) Presence of blocks to reverse transcription:

We analyzed the three intron lariat-exon 2 intermediates for the presence of blocks to reverse transcription in primer extension experiments in which α^{32} P dATP and dideoxynucleotides were used to accurately map the branch points on sequencing gels. As indicated in Figure 5, the intron lariat-exon 2 from *Bam*HI RNA was cleaved in exon 2 by RNase H using oligonucleotide b and further purified by gel electrophoresis prior to reverse

Figure 7. Kinetic analysis of branch point selection on the three human β -globin precursors. Intron lariat-exon 2 from the three precursors were recovered at 20 (lanes a), 40 (lanes b) and 60 min (lanes c) of the splicing reaction (see Materials and Methods) and submitted to primer extension analysis with internal labelling using either deoxyoligonucleotide a (*Ava*II and *Bam*HI RNAs) or deoxyoligonucleotide c (*Acc*I). * Indicates another putative branch.

transcription in order to generate exactly the same template as the intron lariat-exon 2 from *Ava*II RNA. The two RNAs were hybridized to a 16 nt primer (oligonucleotide a) that is complementary to the second exon and the cDNAs synthesized by reverse transcription (Figure 5c and d). Comparison of the sequencing gels reveals that at least three blocks (minor ones are indicated by stars) were consistently present in *Ava*II RNA that map at positions -16, -21 and -36 upstream from the 3' splice site (Figure 5c). These same blocks also appear in *Bam*HI RNA, but one is clearly predominant at position -36 (Figure 5d). Assuming that reverse transcription stops one nucleotide 3' of a branch point, this block exactly maps the branch point to an A residue at 37 nucleotides upstream from the 3' splice site and therefore corresponds to the previously identified branch point (53, 60). Surprisingly, in contrast to the situation seen in *Ava*II and, to a lesser extent, in *Bam*HI RNAs, a similar experiment using as a template the intron lariat-exon 2 from *Acc*I RNA revealed a single block to reverse transcription at position -36 corresponding to the usual branch point (Figure 5a). The same kind of experiment carried out with a 5' labeled primer (oligonucleotide c) confirmed all three stops in *Ava*II RNA

Figure 8. Localization of branch sites by direct RNA analyses. (A) RNase T₁ analysis: Intermediates and splicing products from AvaII and BamHI precursors labeled with $\alpha^{32}\text{P}$ GTP were digested and the products fractionated (Materials and Methods). Lanes 1, 2 and 5, 6 refer to T₁ products from AvaII and BamHI introns lariet, respectively. Similarly, lanes 3, 4 and 7, 8 refer to T₁ products from AvaII and BamHI intron lariet-exon 2, respectively. In lanes 2, 4, 6 and 8, the T₁ products were dephosphorylated prior to electrophoresis. It is noteworthy that dephosphorylation dramatically affects the migration of the smallest T₁ products, so that only the 19 mer can be unambiguously identified. (B–E) RNase A digestion of the modified and normal 10 mer fragments from AvaII intron lariet. The two 10 mer RNase T₁ fragments from the AvaII intron (panel A, lane 1) labeled with $\alpha^{32}\text{P}$ CTP (panels B and D) or $\alpha^{32}\text{P}$ UTP (panels C and E) were withdrawn, digested by RNase A and fractionated by thin-layer chromatography. Panels B and C, then D and E, show the chromatograms corresponding to the normal and modified 10 mer fragments, respectively. Note the presence of the resistant structure X1 in D and E. (F–I) Alkaline hydrolyses of the fast migrating 19 mer and the mixture of 7 mer fragments from AvaII intron lariet. F is a control chromatogram showing the position of the four ribonucleotides phosphates obtained by alkaline hydrolysis of the BamHI precursor labeled with $\alpha^{32}\text{P}$ UTP. H shows the result of an alkaline hydrolysis of the fast migrating 19 mer RNase T₁ fragment labeled with GTP and generated from the AvaII intron (panel A, lane 1). G and I refer to alkaline hydrolyses of the mixture of dephosphorylated 7 mer fragments from AvaII intron

(Figure 6A, lane 2). In control experiments, no blocks to reverse transcription were observed when the *Ava*II RNA precursor was mapped using this same technique (Figure 5b). Furthermore, an intron lariat-exon 2 from *Ava*II RNA was debranched before primer extension analysis. A 154 nucleotides full-length cDNA is synthesized that does not contain blocks to reverse transcription (Figure 6B, lane 1).

b) Direct RNA analysis and localization of the branch structures: Taken together, the above experiments suggest that the two additional blocks observed in *Ava*II RNA correspond to authentic branch points at U residues located at -17 and -22 positions upstream from the 3' splice site. Since the primer extension mapping is indirect, we performed RNA analyses to examine further the nature of the blocks to reverse transcriptase.

In a first experiment we compared electrophoretic mobilities of RNase T₁ fragments from isolated intron lariat products and intron lariat-exon 2 intermediates generated by splicing of *Ava*II and BamHI RNAs labeled with $\alpha^{32}\text{P}$ GTP (Figure 8A). If the blocks to reverse transcription seen in *Ava*II RNA correspond to authentic branch points, then they should change the mobilities of the 10, 7 and 19 nt long fragments containing the -37 , -22 and -17 positions, respectively. This experiment led to the following observations and conclusions: i) The presence of a 11 mer, instead of a 10 mer, among the RNase T₁ products from the excised BamHI intron (lane 5) demonstrates that the branch point at -37 position is effectively present. ii) The presence of a 10 mer in addition to the 11 mer in the case of the *Ava*II intron (lane 1) means that at least another branch point has occurred. iii) The 19 mer region has two bands with the *Ava*II (lane 1) and one with the BamHI (lane 5) introns. Again, this situation implies that at least two branch sites have been used in *Ava*II RNA. Considering that an excised intron has a 3'OH end (see in lane 5 the presence of one 19 mer) as well as the finding that only one 19 mer is present upon alkaline phosphatase digestion of the RNase T₁ products generated from the excised *Ava*II intron (lane 2), we conclude that the fast migrating band contains a branch. It is noteworthy as well that the 19 mer RNase T₁ products from *Ava*II and BamHI intron lariats-exon 2 intermediates have a 3' phosphate end and, therefore, cannot be separated from another band that already contains an additional phosphate (lanes 3 and 7). iv) No conclusion can be drawn from this experiment about the presence of a branch in the 7 nucleotides long fragment.

To further analyse the branch structure contained in the 10 mer fragment derived from the *Ava*II excised intron, we performed experiments as that shown in Figure 8A (lane 1) using RNAs labeled either with ^{32}P CTP (Figure 8B, D) or ^{32}P UTP (Figure 8C, E). The two bands seen within the 10–11 mer region were excised and submitted to RNase A digestion whose products were fractionated by two-dimensional thin-layer chromatography (56). A comparison of the RNase A digestion products of the two T₁ fragments is shown in Figure 8B–E. As a matter of fact, the chromatograms are strictly comparable to those previously described (53, 54) allowing us to conclude that the T₁ fragment ACUCUCUCUG migrating as a 11 mer contains a branch located on the A residue.

To characterize the structure contained in the 19 mer oligonucleotide, we prepared a fast migrating T₁ fragment labeled with ^{32}P GTP and performed alkaline hydrolyses allowing to obtain a mixture of 2' and 3' phosphate mononucleotides which were separated by thin-layer chromatography (Figure 8H). As a control, we carried out the hydrolysis of a BamHI pre-mRNA labeled with ^{32}P UTP (Figure 8F). Clearly, the chromatogram in Figure 8H shows that the label has been transferred to an A residue, as expected from the sequence, as well as to a second structure designated X2 that necessarily corresponds to a branch. X2 is labeled by all but ^{32}P CTP nucleotides (not shown). To identify the branched residue, we used a 19 mer oligonucleotide labeled either with ^{32}P ATP (Figure 8J, L) or ^{32}P UTP (Figure 8K, M). The fragment was dephosphorylated to remove the terminal 3' phosphate, digested with RNase P1 and the products analyzed by thin-layer chromatography before (Figure 8J, K) and after (Figure 8L, M) a second dephosphorylation. Upon dephosphorylation, X2 disappears when the label is on U and, in contrast, persists when the label is on A residues, thus demonstrating that the branch has occurred on a U residue. Taking into account that X2 contains an A residue (Figure 8J, L) only the -17 and -3 U nucleotides appear as possible branch acceptors. If the branch occurred at -3 position, then the A residue at -2 position should not be present as free labeled mononucleotide in Figure 8H, but rather included into X2. Therefore, we conclude that the branch is at -17 position as suggested by the reverse transcription experiments.

To characterize the putative third branch from the excised *Ava*II intron, we analyzed the mixture of RNase T₁ products migrating as a 7 mer oligonucleotides labeled either with all triphosphates (Figure 8I) or with a mixture lacking GTP (Figure 8G). The fragments were dephosphorylated submitted to an alkaline hydrolysis and analyzed by thin-layer chromatography. Figure 8I shows that a resistant structure designated X3 has occurred. It is only labeled when ^{32}P GTP is present implying that the branched nucleotide is the penultimate residue of a T₁ fragment. As our reverse transcription experiments showed a block at -22 position, it appears that the branch found into the mixture of 7 mer oligonucleotides most likely is located on the last U of the CCUAUUG fragment.

*c) The additional branch points in *Ava*II RNA are functional:*

To establish the functionality of these additional branch points, we looked for them in the excised intron lariat. To this end, an excised intron lariat from *Ava*II RNA was therefore purified and submitted to primer extension analysis using oligonucleotide c complementary to the 3' end of the intron (Figure 5e). Under these conditions, -37 and -22 branch points were clearly detected. The -17 branch could not be detected in this experiment as there was no U in the RNA between the end of the primer and the putative branch to direct the incorporation of a radioactive A residue. However, the same kind of experiment carried out with a 5' labeled primer revealed all three stops, including the one at -17 in both the intron lariat-exon 2 and the excised intron lariat (Figure 6A). Taken together these results establish that all three branch sites are functional. In any case, the absence of exon ligation would imply that free exon 2 is a

(Panel A, lane 1). RNA was labeled by all four triphosphates in the experiment 1, while $\alpha^{32}\text{P}$ GTP was omitted in G. (J–M) RNase P1 digestion of the dephosphorylated fast migrating 19 mer from *Ava*II intron lariat. Labeling was with $\alpha^{32}\text{P}$ ATP (J and L) or $\alpha^{32}\text{P}$ UTP (K and M). The RNase P1 products were chromatographed either directly (J and K) or after a subsequent dephosphorylation (L and M). Alkali and RNase P1 cleavage sites as well as the different branch structures and labelling are indicated in the bottom of the Figure.

product of the reaction. Therefore, the relative inefficiency of the second step of the splicing reaction cannot be dependent upon the formation of unusual lariats during the first step. It is noteworthy as well that AccI RNA yields only the usual branch point and, nevertheless, is inefficient.

Is the change of the protection surrounding the branch-3' splice site in BamHI RNA directly related to the selection of branch point?

Intron lariat-exon 2 species from the three RNAs were purified from gels at 20, 40 and 60 min and submitted to primer extension analyses with internal labelling using either a (AvaII and BamHI RNAs) or c (AccI RNA) oligonucleotides as DNA primers (diagrammed in Figure 3). Only one branch site was used in AccI RNA at any times (Figure 7, lanes a,b,c, under AccI). The two other RNA species revealed two differences: i) The usual branch point never becomes predominant in AvaII RNA as it does in BamHI. As a matter of fact, at 20 min, i.e. approximately when covalent modifications begin, all three branch points are used at comparable frequencies in BamHI RNA (Figure 7 BamHI, lane a) while the two additional ones are prevalent in AvaII RNA (Figure 7 AvaII, lane a). Later on, while the three branch points are still used equally in AvaII RNA (Figure 7, lanes b and c under AvaII), the usual one at -37 is clearly prevalent in BamHI RNA (Figure 7 BamHI lanes b and c). ii) The abundance of all lariats decreased in AvaII RNA at 60 min suggesting that extensive debranching occurred before completion of the second step. Similarly this could explain why intron lariat-exon 2 of AccI RNA does not accumulate significantly at 40 and 60 min (Figure 7 AccI lanes b and c). Indeed, debranched intron lariat-exon 2 can be directly detected in AccI RNA where it is easily resolved from exon 1 (Figure 1B, lane d₁) and, with AvaII RNA, by reverse transcription of a band similar to c₂d₂ (Figure 1A, AvaII) which contains unresolved exon 1, intron lariat-exon 2 and debranched intron-exon 2 (data not shown).

DISCUSSION

In this work, we assumed that the relative inability to support the second step of splicing of two human β -globin transcripts truncated at a short distance downstream from the 3' splice site (AvaII and AccI RNAs) may be directly related to an RNP configuration of the corresponding spliceosomes somewhat different from that of a precursor normally efficient in both steps (BamHI RNA).

In a first series of experiments, we used RNase T₁ protection-immunoprecipitation assays (49) with anti-(U2)RNP and anti-Sm antibodies to monitor the snRNPs bound to the branch-3' splice site region of the three RNAs which could be involved in such putative conformational differences. We found that the selection of the branch-3' splice site sequences results from a contact between U2 snRNP and another factor in all but AccI RNAs. In this respect the result agrees with other findings demonstrating that a functional 3' splice site sequence is required for U2 snRNP binding (20, 41, 49). Several components have been described that could bind to the 3' splice site: U5 snRNP (20) via an intron binding protein (37, 38), hnRNP proteins (45, 46) and an U2 auxiliary protein factor termed U2AF (41). Moreover, recent results from our laboratory (48) highly suggested that IBP-U5 snRNP is required at an early stage of the spliceosome assembly

pathway. Since the 5' end of U2 snRNA is required in splicing (17), the most likely possibility is that U2 snRNP interacts with the factor(s) bound to the (Py)_n-3' splice site through RNA-protein contacts. Following this contact, the 3' end of the intron appears less protected in a fully efficient substrate (BamHI RNA). Most likely, the interaction undergoes some modification (possibly a conformational change) allowing the second step of the reaction to proceed. The lack of contact between U2 and the factor(s) recognizing the 3' splice site in AccI RNA, as well as the lack of modification of this contact in AvaII RNA, whose 3' end remains highly protected, could render the second step of the splicing reaction with these precursors less efficient.

In a second series of experiments, careful analysis of branch points revealed two branches on U residues in addition to the usual one on the -37 A whose relative use differed with the RNA substrate and the duration of the reaction: i) AvaII RNA uses all three branch points at all times; ii) BamHI RNA exhibits a similar behaviour at early times but later on shows a strong preference for the usual -37 branch; iii) AccI RNA uses only the -37 A branch at all times.

We can think of three different scenarios to accommodate the above results. The first one implies that the same RNA substrate has the ability to form three structurally different spliceosomes committed to each branch with equal probability. The later accumulation of spliceosomes containing the normal branch on BamHI RNA would call for complex kinetic considerations involving either an increase in the rate of formation or a slower rate of the second step. Under this latter hypothesis, the apparent preference for the -37 branch would simply reflect its poor ability to undergo the ligation step, a consequence which would certainly not be easily acceptable. The second model holds that an homogenous population of initial spliceosomes would sequentially adopt several configurations corresponding to each type of branch. This model is compatible with the shift in protected sequence on BamHI RNA but not with the finding that the three branches are used at equivalent frequencies at all times on AvaII RNA. In the third model, any given spliceosome would undergo a structural transition from an initial undifferentiated state compatible with the three branches to a more rigid configuration able to sustain only the -37 A branching. This transition, reflected by a change in the contacts between U2 snRNP and the 3' factor(s), is indeed observed with BamHI RNA. On the contrary, no such changes are observed with AvaII RNA spliceosome which appears frozen in the undifferentiated configuration. The selection of a unique branch point at the -37 position in AccI RNA could proceed through a somewhat different mechanism. Possibly, a very short exon 2 could render the branch point region of this RNA directly accessible to the U2 snRNP without the need for contacting the 3' factor(s).

The fact that unusual branches are found in the excised lariat has two interesting implications: i) in the context of the latter model, the undifferentiated spliceosome configuration is already functional; ii) a normal branch is not necessary for the ligation step to occur. Neither is it sufficient since the normally branched AccI RNA is spliced with very poor efficiency. This agrees with previously reported results in different systems showing that: i) cryptic branch point activation allows accurate *in vitro* splicing of human and rabbit β -globin intron mutant (32, 53); ii) alternative branch points are selected during splicing of a yeast pre-mRNA in mammalian and yeast extracts (61); iii) several branch points can occur, possibly on C or U residues (62-64).

ACKNOWLEDGEMENTS

Thanks are due to Ch. Aymard and R. Dietz for excellent technical assistance, to Drs. T. Maniatis, J. Steitz and S. Hoch for the respective gifts of human β -globin plasmids, anti-(U2) RNP and anti-Sm 7.13 antibodies. This work was supported by grants from : CNRS (URA 1191), INSERM (contrat n.871001), ARC, the Ligue Nationale contre le Cancer, and the Fondation pour la Recherche Médicale. J. Tazi, J. Temsamani and C. Alibert were supported by fellowships from the Association pour la Recherche contre le Cancer and the Ligue Régionale contre le Cancer (comité de l'Hérault) and the Ligue Nationale contre le Cancer, respectively.

REFERENCES

- Bindereif, A., and Green, M.R. (1986) *Mol. Cell. Biol.* 6, 2582–2592.
- Bindereif, A., and Green, M.R. (1987) *EMBO J.* 6, 2415–2424.
- Brody, E., and Abelson, J. (1985) *Science* 228, 963–967.
- Frendewey, D., and Keller, W. (1985) *Cell* 42, 355–367.
- Grabowski, P.J., Seiler, S. and Sharp, P.A. (1985) *Cell* 42, 345–353.
- Perkins, K.K., Furneaux, H.M., and Hurwitz, J. (1986) *Proc. Natl. Acad. Sci. USA* 83, 887–891.
- Brunel, C., Sri-Widada, J., and Jeanteur, Ph. (1985) *Prog. Mol. Subcell. Biol.* (Hahn, F.E., Kopecko, D.J. & Müller, W.E.G., eds), vol. 9, pp. 1–52, Springer-Verlag, Berlin.
- Green, M.R. (1986) *Ann. Rev. Genet.* 20, 671–708.
- Maniatis, T., and Reed, R. (1987) *Nature* 325, 673–678.
- Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986) *Ann. Rev. Biochem.* 55, 119–150.
- Sharp, P.A. (1987) *Science* 235, 766–771.
- Bozzoni, I., Annesi, F., Beccari, E., Frapapan, P., Pierandrei-Amaldi, P., and Amaldi, F. (1984) *J. Mol. Biol.* 180, 1173–1178.
- Padgett, R., Mount, S., Steitz, J., and Sharp, P.A. (1983) *Cell* 35, 101–107.
- Yang, V.W., Lerner, M.R., Steitz, J.A. and Flint, S.J. (1981) *Proc. Natl. Acad. Sci. USA* 78, 1371–1375.
- Berget, S.M., and Roberson, B.L. (1986) *Cell* 46, 691–696.
- Black, D.L., Chabot, B., and Steitz, J.A. (1985) *Cell* 42, 737–750.
- Black, D.L., and Steitz, J.A. (1986) *Cell* 46, 697–704.
- Krainer, A.R., and Maniatis, T. (1985) *Cell* 42, 725–736.
- Krämer, A., Keller, W., Appel, B., and Lührmann, R. (1984) *Cell* 38, 299–307.
- Chabot, B., Black, D.L., LeMaster, D.M., and Steitz, J.A. (1985) *Science* 230, 1344–1349.
- Sri-widada, J., Liautard, J.P., Brunel, C., and Jeanteur, Ph. (1983) *Nucl. Acids Res.* 11, 6631–6646.
- Grabowski, P.J. and Sharp, P.A. (1986) *Science* 233, 1294–1299.
- Konarska, M.M. and Sharp, P.A. (1986) *Cell* 46, 845–855.
- Konarska, M.M. and Sharp, P.A. (1987) *Cell* 49, 763–774.
- Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J.A. (1983) *Cell* 33, 509–518.
- Ruskin, B., and Green, M.R. (1985) *Science* 229, 135–140.
- Krämer, A. (1987) *J. Mol. Biol.* 196, 559–573.
- Séraphin, B., Kretzner, L. and Rosbash, M. (1987) *EMBO J.* 7, 2533–2538.
- Zhuang, Y., and Weiner, A.M. (1986) *Cell* 46, 827–835.
- Parker, R., Siliciano, P.G., and Guthrie, C. (1987) *Cell* 49, 229–239.
- Padgett, R.A., Konarska, M.M., Aebi, M., Hornig, H., Weissmann, C., and Sharp, P.A. (1985) *Proc. Natl. Acad. Sci. USA* 82, 8349–8353.
- Ruskin, B., Greene, J.M., and Green, M.R. (1985) *Cell* 41, 833–844.
- Reed, R., and Maniatis, T. (1988) *Genes Dev.* 2, 1268–1276.
- Zhuang, Y., Goldstein, M. and Weiner, A.M. (1989) *Proc. Natl. Acad. Sci. USA* 86, 2752–2756.
- Zhuang, Y. and Weiner, A.M. (1989) *Genes Dev.* 3, 1545–1552.
- Wu, J., and Manley, J.L. (1989) *Genes Dev.* 3, 1553–1561.
- Tazi, J., Alibert, C., Temsamani, J., Reveillaud, I., Cathala, G. Brunel, C., and Jeanteur, Ph. (1986) *Cell* 47, 755–766.
- Gerke, V., and Steitz, J.A. (1986) *Cell* 47, 973–984.
- Lamond, A.I., Konarska, M.M., Grabowski, P.J., and Sharp, P.A. (1988) *Proc. Natl. Acad. Sci. USA*, 85, 411–415.
- Losski, M., Anderson, G., Jackson, S., Beggs, J. (1987) *Cell* 51, 1019–1026.
- Ruskin, B., Zamore, P.D., and Green, M.R. (1988) *Cell* 52, 207–219.
- Ruby, S.W., and Abelson, J. (1988) *Science* 242, 1028–1035.
- Zillman, M., Rose, S. and Berget, S. (1987) *Mol. Cell. Biol.* 7, 2877–2883.
- Séraphin, B., and Rosbash, M. (1989) *Cell* 59, 349–358.
- Choi, Y. D., Grabowski, P.J., Sharp, P.A. and Dreyfuss, G. (1986) *Science* 231, 1534–1539.
- Swanson, M.S., and Dreyfuss, G. (1988) *EMBO J.*, 7, 3519–3529.
- Nelson, K.K., and Green, M.R. (1989) *Genes Dev.* 3, 1562–1571.
- Tazi, J., Temsamani, J., Alibert, C., Rhead, W., Khellil, S., Cathala, G. Brunel, C., and Jeanteur, Ph. (1989) *Nucl. Acids Res.* 17, 5223–5243.
- Chabot, B., and Steitz, J.A. (1987) *Mol. Cell. Biol.* 7, 281–293.
- Krainer, A., Maniatis, T., Ruskin, B., and Green, M. (1984) *Cell* 36, 993–1005.
- Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) *Nucl. Acids Res.* 11, 1475–1489.
- Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular Cloning: a laboratory manual* (Cold Spring Harbor, New York: CSH Laboratory).
- Ruskin, B., Krainer, A., Maniatis, T., and Green, M. (1984) *Cell* 38, 317–331.
- Konarska, M.M., Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1985) *Nature* 313, 552–557.
- Wallace, J. and Edmonds M. (1983) *Proc. Natl. Acad. Sci. USA* 80, 950–954.
- Silberklang, M., Gillum, A.M., and RajBhandary, U.L. (1979) *Meth. Enzymol.* 59, 58–109.
- Furdon, P., and Koe, R. (1986) *Proc. Natl. Acad. Sci. USA* 83, 927–931.
- Reed, R., and Maniatis, T. (1986) *Cell* 46, 681–690.
- Lawn, R.M., Efstratiadis, A., O'Connell, C., and Maniatis, T. (1980) *Cell* 21, 647–651.
- Reed, R., and Maniatis, T. (1985) *Cell* 41, 95–105.
- Ruskin, B., Pikielny, C.W., Rosbash, M., and Green, M.R. (1986) *Proc. Natl. Acad. Sci. USA* 83, 2022–2026.
- Gattoni, R., Schmitt, P., and Stevenin, J. (1988) *Nucl. Acids. Res.* 16, 2389–2409.
- Harmuth, K. and Barta, A. (1988) *Mol. Cell. Biol.* 8, 2011–2020.
- Noble, J.C.S., Zhen-Qiang Pan, Prives, C. and Manley, J.L. (1987) *Cell* 50, 227–236.