Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 704–708, February 1988
Biochemistry

Structure of a pre-mRNA branch point/3′ splice site region

(NMR/nuclease mapping/RNA secondary structure)

KATHLEEN B. HALL*, MICHAEL R. GREEN†‡, AND ALFRED G. REDFIELD*‡

*Biochemistry Department, Brandeis University, Waltham, MA 02254; and †Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138

Contributed by Alfred G. Redfield, September 28, 1987

ABSTRACT **We have analyzed the solution structure of RNA containing the branch point/3′ splice site region of the first intervening sequence (IVS*1*) of human β-globin pre-mRNA by nuclease mapping and NMR. Nuclease mapping indicates that there are two distinct structural domains: one contains the branch point region, and the other includes the 3′ splice site and second exon sequences. In the branch point domain, the adenosine at which the RNA branch forms appears to be in the loop of a stem/loop structure. The branch point structure does not appear to interact with other parts of the RNA, since its unique nuclease digestion pattern is conserved among transcripts containing the entire intron or only the branch point region. This is confirmed by a comparison of the NMR spectra of two RNA transcripts; a distinct set of resonances appears in the spectra of the RNA containing only branch sequences or including 3′ splice site/exon 2 sequences. NMR studies further show that the 3′ splice site/exon 2 domain has a lower melting temperature than the branch point domain, suggesting that the two regions are distinct dynamically as well as structurally. Nuclease mapping studies of adenovirus major late IVS*1* indicate that this RNA has structural features in common with the human β-globin transcript.**

The primary transcripts (pre-mRNAs) of most structural genes in higher eukaryotes contain intervening sequences (IVS; introns) that are removed by splicing. A large ribonucleoprotein complex mediates the two-step splicing reaction (1–4). In the first stage, the pre-mRNA is cleaved at the 5′ splice site generating two splicing intermediates, the first exon RNA species and an RNA species composed of the intron and second exon (IVS-exon 2 RNA species). In the second stage, cleavage at the 3′ splice site and ligation of the exons occurs, resulting in the excision of the intact intron. The IVS-exon 2 splicing intermediate and the excised IVS are in the form of a lariat in which the 5′ end of the intron is joined to an adenosine residue near the 3′ end of the intron by a 2′-5′ phosphodiester bond.

Although significant progress has been made in understanding the pathway of splicing, a major unsolved problem is the mechanism of accurate splice site selection. Three sequence elements have been shown to participate in the pre-mRNA splicing reaction: the 5′ splice site, the 3′ splice site, and the branch point sequence. The 5′ and 3′ splice sites of mammalian pre-mRNAs conform to reasonably well-defined consensus sequences (5). However, in higher eukaryotes the branch point displays considerable sequence variability and appears to be primarily selected by distance from the 3′ splice site; RNA branch formation occurs at an adenosine residue that lies between 18 and 37 nucleotides (nt) upstream from the 3′ splice site (2).

The three sequence elements seem insufficient to unambiguously account for accurate splicing (1, 2, 6, 7). Other elements that could confer specificity to the splicing process include RNA–protein interactions, RNA·RNA interactions, and the secondary/tertiary structure of the pre-mRNA (1, 2, 8). For example, RNA structure determines splice site selection in class I and class II self-splicing introns (9).

We have used the human β-globin IVS*1* sequence to address the question of whether the RNA structure could be important in metazoan pre-mRNA splicing. In particular, we looked at the branch point/3′ splice site sequences, since this region of the pre-mRNA interacts with protein and small nuclear ribonucleoprotein (snRNP) components of the splicing complex. By using a combination of nuclease mapping and NMR techniques, we have found that there is a distinct structure around the branch point. Specific features of this structure, which appear also in the adenovirus major late (AdML) IVS*1*, may play a role in the splicing process.

MATERIALS AND METHODS

SP6 RNA polymerase, RNasin, and restriction enzymes were from Promega Biotec (Madison, WI). S1 nuclease, cobra venom V1, and RNase T1 were from P/L Biochemicals. RNase U2 was from Calbiochem. T7 RNA polymerase was isolated from the overproducing clone provided by J. Dunn (Brookhaven National Laboratory) (10).

The original clone 5′ΔHX of the human β-globin IVS*1* was described by Ruskin and Green (11). The *Xba* I/*Bam*HI fragment was cloned into pGEM2 under control of the T7 RNA polymerase promoter, to give the new clone T7Δ5′HX1. The truncated clone T7Δ5′HX2 was constructed by removing the *Mst* II/*Bam*HI fragment from 5′ΔHX containing the 3′ splice site and the exon 2 sequences. The fragment of the AdML IVS*1* used for nuclease mapping experiments was constructed by using the *Hin*dIII site located 82 nt upstream of the 3′ splice site and the *Bgl* I site located 12 nt downstream of the 3′ splice site.

For nuclease mapping, the T7- or SP6-derived RNA transcripts were 3′ end-labeled by using cytidine 5′-[^{32}P]-phosphate and T4 RNA ligase and were gel-purified. To obtain partial digestion, the amount of nuclease added was determined by titration using a constant amount of RNA. The following buffers were used: for T1 and cobra venom V1 ribonuclease, 200 mM NaCl/5 mM $MgCl_2$/10 mM Tris, pH 7.4; for S1 nuclease, 200 mM NaCl/5 mM $MgCl_2$/0.1 mM $ZnSO_4$/25 mM sodium acetate, pH 5.5 (12); for U2 nuclease, S1 buffer without the $ZnSO_4$.

NMR samples contained 2 mM RNA in 10 mM sodium phosphate, pH 7/1 mM EDTA/10% ^2H_2O/100 mM NaCl. NMR spectra of the exchangeable imino protons were taken in water by using either a 214 or a JxRy pulse (13). Spectra are reported in ppm from 2,2-dimethyl-2-silapentane-5-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IVS, intervening sequence(s); snRNP, small nuclear ribonucleoprotein; NOE, nuclear Overhauser effect; AdML, adenovirus major late; nt, nucleotide(s); t_m, melting temperature.

‡To whom reprint requests should be addressed.

sulfonate based on H_2O as a reference, assumed to be 4.8 ppm.

RESULTS

The Branch Point Region Forms a Discrete Structure. The secondary structure of the human β-globin IVS*1* branch point/3′ splice site region was probed by nuclease digestion using three RNA substrates. One RNA contains the first exon, the entire IVS*1*, and 21 nt of exon 2 (Δ6HX). A shorter RNA contains the complete branch point region and downstream sequences including 21 nt of exon 2 but lacks the 5′ splice site (T7Δ5′HX1). This intron sequence with about 200 nt of the second exon was shown by RNase A protection experiments to bind factors associated with the branch point (11). The shortest RNA, T7Δ5′HX2, contains the branch point region and the polypyrimidine tract but lacks the 5′ and 3′ splice sites as well as exon 2. The use of these three transcripts allowed us to determine whether long-range interactions contributed to the structure near the branch point.

S1 nuclease, which is specific for any single-stranded RNA, cleaves all three transcripts at the adenosine used for branch formation (position −37, see Fig. 1) and at 4 downstream nucleotides (−36 to −33). There is also a strong S1 nuclease cleavage region in part of the polypyrimidine tract (positions −17 to −11). Positions −24 to −21 are weak S1 nuclease cleavage sites. RNase U2, which cleaves at ApN (N = any nucleotide) sequences in single-stranded RNA, has two major cleavage sites in the RNA transcripts: one at position −37 and the other at nucleotide −16. Weaker cleavage is seen at positions −24 and −2 (the ApG at the 3′ splice site). The single-strand sequence GpN is cleaved by nuclease T1. In these RNAs, there is strong T1 cleavage at −38 (the guanosine adjacent to the adenosine used for branch formation) and weaker sites at −22, −21, and −1, +1 (the 3′ splice site). The double-strand-specific RNase cobra venom V1 cleaves at positions −43 to −39 (immediately upstream of the branch position) and more weakly at positions −30 to −27 (downstream of the branch point). Nucleotides −27 to −25 and nucleotide −19 are also sites of strong cleavage by V1, with other weaker sites at −48 to −45 and −5 to −1 (the 3′ splice site). From these results, we conclude that the adenosine used for branch formation and a portion of the polypyrimidine tract are present as single-stranded regions in the pre-mRNA, independent of the presence of the 5′ splice site, 3′ splice site, or exon 2 sequences.

NMR Spectra Identify Two Domains. NMR provides information that supplements nuclease mapping data. For example, the number of A·U, G·C, and G·U base pairs in the structure can be estimated, and helical regions can be identified. To confirm by NMR that the branch sequences form a unique domain, the two transcripts T7Δ5′HX1 and T7Δ5′HX2 containing the branch point sequences were used. These NMR experiments were done in H_2O, so that the imino protons on the paired bases (NH....N) are visible. Fig. 2 shows the ^1H NMR spectra of the two transcripts. It is clear that the two spectra share a common subset of resonances (peaks a–c, e–i, k, o, and r), which appear at the same positions in both spectra (see Table 1). Since only the branch point sequences (−64 to −4) are present in both transcripts, the peaks must arise from the structure of this region of the RNA. This conclusion is supported by temperature and nuclear Overhauser effect (NOE) studies.

The NMR spectra as a function of temperature show the melting of individual base pairs. Such a temperature study of T7Δ5′HX1 is shown in Fig. 3. At 60°C, most of the peaks are still present in the spectrum, although a loss of intensity indicates that some melting has begun. By 65°C, most of the resonances associated with the 3′ splice site/exon 2 sequences (−3 to +21) have disappeared. The different temperature dependence of the two portions of the transcript again indicates that the two regions of the RNA are forming separate domains.

In a separate experiment, the melting temperature (t_m) of the T7Δ5′HX1 transcript was measured optically (data not shown). In NMR buffer, the t_m of the 2 μM RNA was 65°C, with 19% hyperchromicity. The t_m measured at optical concentrations agrees well with the apparent NMR t_m of the transcript at 2 mM, suggesting that there is no aggregation or bimolecular association in the NMR experiments. Addition of $MgCl_2$ to 5 mM in the 2 μM sample increased the t_m by ≈10°C. Addition of 10 mM $MgCl_2$ to the NMR sample (creating a five times molar excess of Mg^{2+}) causes few spectral changes, the most noticeable one being the stabili-

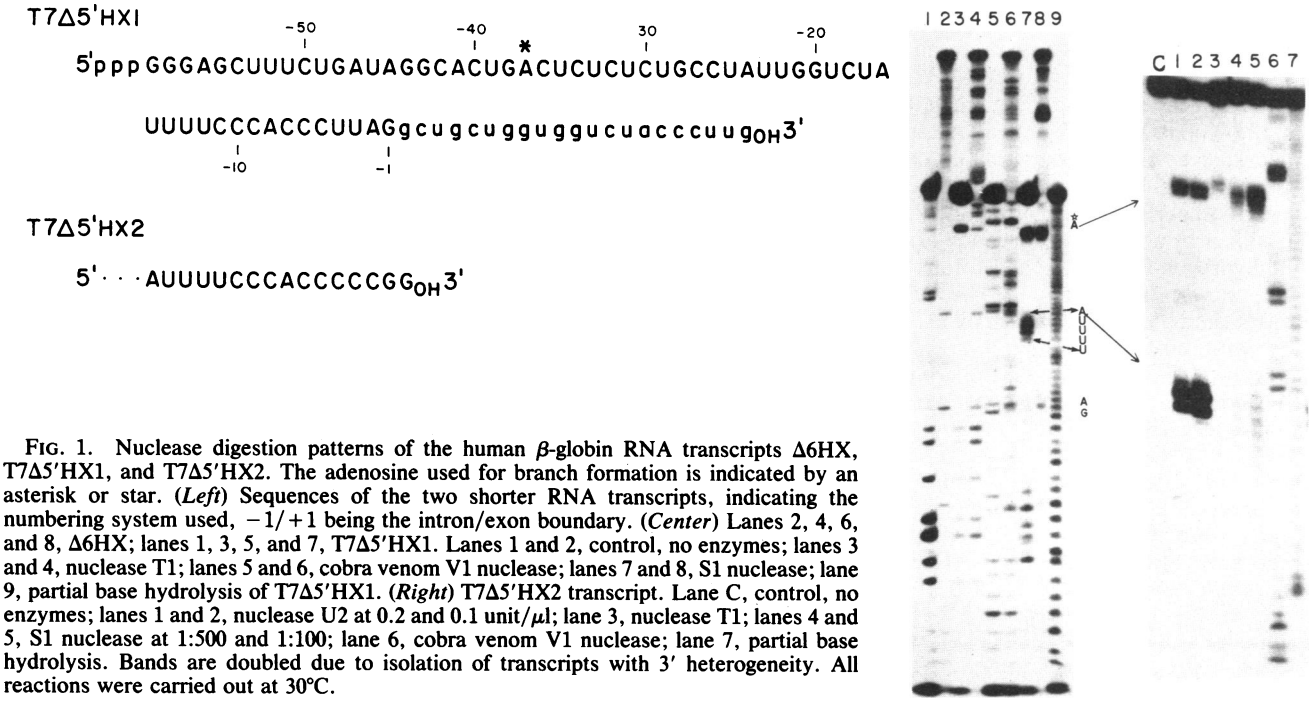


FIG. 1. Nuclease digestion patterns of the human β-globin RNA transcripts Δ6HX, T7Δ5′HX1, and T7Δ5′HX2. The adenosine used for branch formation is indicated by an asterisk or star. (*Left*) Sequences of the two shorter RNA transcripts, indicating the numbering system used, −1/+1 being the intron/exon boundary. (*Center*) Lanes 2, 4, 6, and 8, Δ6HX; lanes 1, 3, 5, and 7, T7Δ5′HX1. Lanes 1 and 2, control, no enzymes; lanes 3 and 4, nuclease T1; lanes 5 and 6, cobra venom V1 nuclease; lanes 7 and 8, S1 nuclease; lane 9, partial base hydrolysis of T7Δ5′HX1. (*Right*) T7Δ5′HX2 transcript. Lane C, control, no enzymes; lanes 1 and 2, nuclease U2 at 0.2 and 0.1 unit/μl; lane 3, nuclease T1; lanes 4 and 5, S1 nuclease at 1:500 and 1:100; lane 6, cobra venom V1 nuclease; lane 7, partial base hydrolysis. Bands are doubled due to isolation of transcripts with 3′ heterogeneity. All reactions were carried out at 30°C.

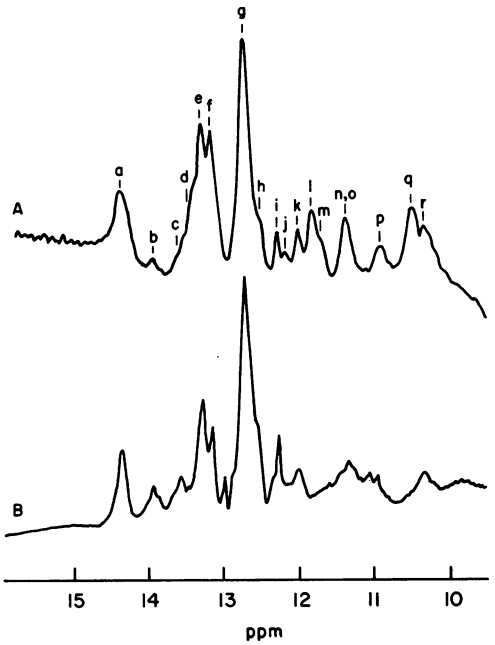


FIG. 2. ^{1}H NMR spectra of the imino protons of T7Δ5′HX1 (*A*) and T7Δ5′HX2 (*B*) in H_2O at 30°C. Spectra were acquired on a 500-MHz spectrometer. Each spectrum consisted of 1600 transients, using 8-kHz spectral width at 1-kHz resolution. Recovery time was 200 msec for T7Δ5′HX1 and 300 msec for T7Δ5′HX2.

zation of 1 G·U pair (peaks o/r). This result suggests that there are no specific associations between base pairs and Mg^{2+} that could in turn cause Mg^{2+}-dependent conformational changes.

The A·U and G·U base pairs in the NMR spectra were identified by NOE experiments by selectively irradiating at the frequency of a specific hydrogen-bonded imino proton and observing its dipolar relaxation to a proton on the same base pair [to either the AH2 or GH1/UH3, respectively (14)]. Based on the results of these experiments, the structure of the branch point region appears to be comprised of 1 G·U base pair (peaks o/r), 4 A·U base pairs (peaks a–c), and at least 8 G·C pairs (peaks e, f, g1, g2, and h–k). The remaining 4 G·U pairs (peaks l/q, m/q, l/p, and n/j) come from the splice site domain (−3 to +21) as do 1 A·U pair (peak d) and several G·C pairs (e and g). There are two peaks in the spectrum of the shorter RNA (T7Δ5′HX2) that are not present in the transcript containing the 3′ splice site (peak f′ at 13.07 ppm and peak i′ at 12.67 ppm). These peaks, both identified as G·C pairs, probably come from a small stem structure formed at the 3′ end of the T7Δ5′HX2 transcript.

NOE experiments can also indicate the order of base pairs in helices by allowing observation of the dipolar relaxation of the hydrogen-bonded imino protons to imino protons of the base pair on either side of the base pair of interest (15). Table 1 lists the NOEs observed between peaks in the spectra of the two transcripts.

From NOE experiments, the branch point region in both transcripts is shown to be comprised of several stacked sequences. There are at least 3 adjacent G·C/G·C pairs, 1 A·U/G·C/G·C sequence, 4 adjacent A·U/G·C pairs, and 1 G·U pair adjacent to a G·C pair. The G·U/G·C NOE is observed below 30°C. Other NOEs from base pairs in this region were observed at 50°C, suggesting that the stacking is relatively stable. NOE experiments on the 3′ splice site/exon 2 sequences of T7Δ5′HX1 indicate that the 7 base pairs in this domain are arranged in two short helices: one sequence of G·C/G·U/A·U/G·U; the other, G·U/G·C/G·U. Some of the small stacked regions identified in the RNAs may be involved in tertiary structures, and others may be

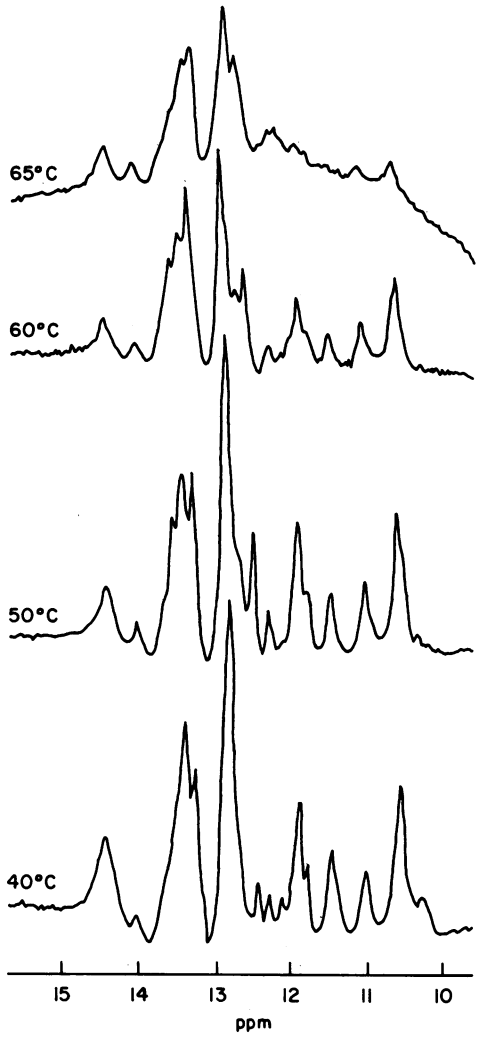


FIG. 3. ^{1}H NMR spectra of the imino protons of the T7Δ5′HX1 transcript as a function of temperature. Each spectrum is the average of 800 transients. At 75°C, little intensity remains, indicating that the molecule has melted. After heating, the sample was cooled, and a spectrum was taken at 30°C to establish that there was no degradation.

part of longer helices. The presence of these stacked segments is used as a set of constraints to help produce a structural model (see *Discussion*).

Nuclease Mapping the AdML IVS*1*. To determine whether another pre-mRNA contains structural features comparable to human β-globin IVS*1*, we analyzed the AdML IVS*1* by nuclease digestion. This pre-mRNA was used because, like human β-globin, it is efficiently spliced *in vitro* (16).

In the AdML pre-mRNA there are strong S1 nuclease cleavages from −15 to −7 in the polypyrimidine tract (Fig. 4) and from −70 to −65, with weak cleavages from −39 to −35. There are major RNase U2 cleavages at positions −38 and −35 and minor sites at −28, −24 (the adenosine used for branch formation) and −2 (the 3′ splice site ApG). RNase T1 cleaves at positions −36, −65, and −1 (the 3′ splice site) and with lower efficiency at position −19. RNase V1 cleavage is strongest at positions −17 to −15 just upstream of the polypyrimidine tract and at −40 and −50, with weaker sites from −29 to −27.

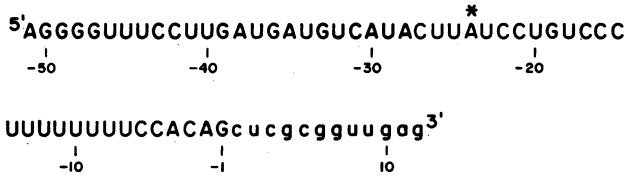
A comparison of the nuclease mapping patterns of the human β-globin and AdML pre-mRNAs shows some similarities. The sequences around the adenosine used for branch formation are single-stranded, although in AdML there is some ambiguity as to the extent of the single-

Table 1. Summary of NOEs observed

Resonance	Base pair type	T7Δ5'HX1	T7Δ5'HX2	Comments
a	A·U	e, h	e, h	NOEs to 7.84, 7.44
b	A·U	c, g	c, f, g	NOEs to 7.62, 7.42
c	A·U	NO	NO	NOEs to 7.39, 7.07
d	A·U	n, j, l*, q*	NP	NOE to 7.67
e	G·C	a, g, l, p, q	a, b, g, o	
f	G·C	i, g	i, g	
g	G·C	b, e, f, g, n	f	
h	G·C	NO	NO	
i	G·C	f, g	f	
j	G·U	n, d	NP	Strong NOE to n
k	G·C	NO	c	
l	G·U	e, p, q	NP	Strong NOEs to p, q
m	G·U	e, q	NP	Strong NOE to q
n	G·U	d, j, g	NP	Strong NOE to j
o	G·U	r	r, e	Temperature and Mg^{2+} sensitive
p	G·U	l	NP	Strong NOE to l
q	G·U	e, l, m	NP	Strong NOEs to l, m
r	G·U	o	o, g	Temperature and Mg^{2+} sensitive

NO, none observed; NP, not present.
*Spillover from e.

stranded region (no S1 cleavage at −24). One region of the polypyrimidine tract is single-stranded in both, a stretch of UAUUUUC at positions −17 to −11 in human β-globin and a stretch of CUUUUUUUU at positions −15 to −7 in the AdML RNA. The ApG at the 3′ splice sites is a weak



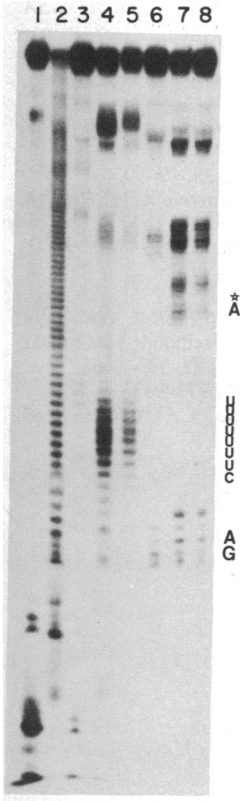


FIG. 4. (*Upper*) Sequence of the AdML RNA transcript used for nuclease mapping experiments. The −1/+1 indicates the 3′ intron/exon boundary, and the adenosine at −24 is used for branch formation. (*Lower*) Nuclease mapping of the AdML IVS*1* transcript. Lane 1, control, no enzyme; lane 2, partial base hydrolysis; lane 3, cobra venom V1 nuclease; lanes 4 and 5, S1 nuclease at 1:100 and 1:500; lane 6, nuclease T1; lanes 7 and 8, nuclease U2 at 1:10 and 1:20. The adenosine used for branch formation is indicated by an asterisk or star.

cleavage site in both RNAs, perhaps because it is involved in some tertiary interaction. The structural similarities between these two RNAs suggest that they may have a common structure around the branch point region even though the sequences are not conserved.

DISCUSSION

Based on nuclease mapping and NMR data, we propose a structural model of the 85-nt human β-globin IVS*1* fragment that contains two noninteracting domains (Fig. 5). The first domain includes the sequences around the branch point and the polypyrimidine tract. A stem/loop structure can be constructed in which the adenosine used for branch formation is in the loop, accounting for its sensitivity to U2 and S1 nuclease digestion. One part of the polypyrimidine tract is also single-stranded, and again it is possible to construct a stem/loop structure, although the stem is very short. The nuclease sensitivity of the UAUUUU sequence of the polypyrimidine tract is easily explained by this stem/loop structure.

The base pairs in the stems of the model agree moderately well with NOE data. The sequence A·U/G·C/G·C in the branch domain found by NOE experiments appears three

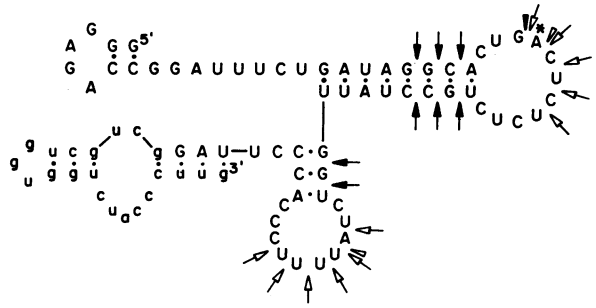


FIG. 5. Structural model of the branch point/3′ splice site portion of the human β-globin IVS*1*. Intron sequences are in uppercase letters; exon 2 sequences in lowercase letters. Sites of primary nuclease cleavages are marked: V1 by closed, tailed arrows; S1 by open, tailed arrows; U2 by open arrowheads; and T1 by a closed arrowhead. The adenosine used for branch formation is indicated by an asterisk.

times in the proposed structure. If it is assigned to the short helix within the polypyrimidine tract, then the longer branch stem must account for the other stacked segments found by NMR. Alternatively, that A·U/G·C/G·C trimer could be located in the branch stem, either adjacent to the loop or internally in the stem. In either case, the other stem regions can account for all of the G·C/G·C stacks found by NOEs and for two of the four G·C/A·U stacks identified by NMR. Those observed NOEs that do not appear in the model may be due to tertiary interactions not included in the secondary structure. In addition, the stems in the model predict two A·U/A·U stacks that so far have not been seen in NOE experiments. The 1 G·U pair in the branch domain is found in the model at the end of a helix. NMR experiments show that this base pair (peaks o/r) is temperature labile and is stabilized by Mg^{2+}, both properties that would be expected from a terminal base pair. The second domain includes the 3′ splice site/exon 2 sequence and is found in the T7Δ5′HX1 transcript. Combining the NMR and nuclease mapping data produces a structure with the two helices found by NOEs separated by an unpaired bulge. The ApG dinucleotide at the 3′ splice site is base-paired, but because it is near the end of the helix, it is likely to fray and therefore would be sensitive to cleavage by all nucleases. The 4 G·U base pairs found by NMR are all accounted for in this proposed structure.

Flexibility of the 3′ splice site may be a feature of many pre-mRNAs. Earlier experiments by Munroe and Duthie (17) using several adenovirus introns flanked by complete second exons also show examples of cleavage at the splice site ApG by single- and double-strand-specific nucleases, indicating that the RNA structure may be dynamic (or involved in tertiary structure). These results are important for comparison with the data reported here, since the lack of a complete exon 2 might result in an abnormal structure at the 3′ splice site.

Because the structure formed by the RNA sequence of the branch point region of the human β-globin IVS*1* is distinctive and stable, it seems reasonable to propose that it may play a role in splicing. In the normal pre-mRNAs analyzed in our experiments, the nucleotides around the adenosine used for branch formation were unpaired and presumably in a loop. When the structure of the branch point region of eight cryptic mutants was enzymatically probed, the accessibility of the adenosine used for branch formation to S1 and U2 nucleases was found to vary (data not shown). The splicing efficiency of many of these RNAs is reduced *in vitro* relative to normal efficiency (18), which could reflect, in part, structural differences around the branch point. This would suggest that accessibility of the adenosine used for branch formation and of surrounding nucleotides may be necessary for optimal use of the pre-mRNA by the splicing apparatus.

These studies show that in solution, the RNA adopts a distinct structure. However, it is not clear if this structure is maintained in the presence of the associated proteins and snRNPs of the splicing complex or what its function would be. Solnick (19) demonstrated that a long stem of 105 nt was maintained *in vitro* and in *in vivo* splicing conditions. It must be noted, though, that *in vitro* effects are often not apparent *in vivo*, where smaller stems (<50 base pairs) seem to present no obstacle to accurate splicing (20).

This region of the pre-mRNA is implicated in interactions with the factors U2 snRNP (1–4) and a 70- to 100-kDa protein (21, 22). If there are specific pre-mRNA sequences that participate in these interactions, such as base pairing with the U2 snRNA in analogy with the *Saccharomyces cerevisiae* UACUAAC box (23), then the structure of the pre-mRNA may provide a mechanism to make those sequences accessible, such as by sequestering them in a loop. Thus, the RNA structure could function as a recognition element or could influence the kinetics or efficiency of splicing. By selectively altering structural features of the pre-mRNA and then observing specific interactions with various components of the splicing complex, we hope to determine the features that control accurate and efficient splicing.

This work was supported by grants from the National Institutes of Health to A.G.R. and M.R.G. and by a Damon Runyon-Walter Winchell Cancer Fund Fellowship to K.B.H. This is publication no. 1632 from the Brandeis University Biochemistry Department.

1. Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiler, S. & Sharp, P. A. (1986) *Annu. Rev. Biochem.* **55**, 1119–1150.
2. Green, M. R. (1986) *Annu. Rev. Genet.* **20**, 671–708.
3. Sharp, P. A. (1987) *Science* **235**, 766–771.
4. Maniatis, T. & Reed, R. (1987) *Nature (London)* **325**, 673–687.
5. Mount, S. M. (1982) *Nucleic Acids Res.* **10**, 459–472.
6. Wieringa, B., Meyer, F., Reiser, J. & Weissman, C. (1983) *Nature (London)* **301**, 38–43.
7. Reed, R. & Maniatis, T. (1986) *Cell* **46**, 681–690.
8. Kuhne, T., Wieringa, B., Reiser, J. & Weissman, C. (1983) *EMBO J.* **2**, 727–733.
9. Cech, T. R. & Bass, B. L. (1986) *Annu. Rev. Biochem.* **55**, 599–629.
10. Davanloo, P., Rosenberg, A. H., Dunn, J. J. & Studier, F. W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2035–2039.
11. Ruskin, B. & Green, M. R. (1985) *Cell* **43**, 131–142.
12. Cech, T. R., Tanner, N. K., Tinoco, I., Jr., Weir, B. R., Zuker, M. & Perlman, P. S. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3903–3907.
13. Redfield, A. G. (1983) *Chem. Phys. Lett.* **96**, 537–540.
14. Johnston, P. D. & Redfield, A. G. (1978) *Nucleic Acids Res.* **5**, 3913–3927.
15. Roy, S. & Redfield, A. G. (1981) *Nucleic Acids Res.* **9**, 7073–7083.
16. Hernandez, N. & Keller, W. (1983) *Cell* **35**, 89–99.
17. Munroe, S. H. & Duthie, R. S. (1986) *Nucleic Acids Res.* **14**, 8447–8465.
18. Ruskin, B., Greene, J. & Green, M. R. (1985) *Cell* **41**, 833–844.
19. Solnick, D. (1985) *Cell* **43**, 667–676.
20. Solnick, D. & Lee, S. I. (1987) *Mol. Cell. Biol.* **7**, 3194–3198.
21. Tazi, J., Alibert, C., Temsamani, J., Reveillaud, I., Cathala, G., Brunel, C. & Jeanteur, P. (1986) *Cell* **47**, 755–766.
22. Gerke, V. & Steitz, J. A. (1986) *Cell* **47**, 973–984.
23. Parker, R., Siliciano, P. G. & Guthrie, C. (1987) *Cell* **49**, 229–239.