BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **237,** 182–187 (1997)
ARTICLE NO. RC977091

The D1-A2 and D2-A2 Pairs of Splice Sites from Human Immunodeficiency Virus Type 1 Are Highly Efficient *in Vitro,* in Spite of an Unusual Branch Site

Laurence Damier, Lionel Domenjoud,[1] and Christiane Branlant[2]
Laboratoire d'Enzymologie et de Génie Génétique, URA CNRS 457, Université Henri Poincaré Nancy 1, Boulevard des Aiguillettes, BP239, 54506 Vandoeuvre les Nancy Cedex, France

Received June 24, 1997

Using *in vitro* splicing assays with HeLa cell nuclear extracts, we showed that the HIV-1 pairs of splice sites D1-A2 and D2-A2 are efficiently used *in vitro,* as compared to the control D1-A2 pair of sites from the E3 transcription unit of human adenovirus-2. The strong efficiency of the two HIV-1 pairs of sites is surprising, as we also showed by primer extension analysis that the branch-site sequence used at the HIV-1 acceptor site A2 is UAGCAGA, with a dominant utilization of the ultimate G as the branched residue. No significant increase of the splicing efficiency was observed upon replacement of the wild-type branch-site sequence by a canonical sequence, in spite of the utilization of an A residue as the branched nucleotide. Results are discussed taking into account the present knowledge on branch-site selection.

Retroviruses transcribe their RNA from an integrated proviral genome and utilize the host cell machinery to produce their mRNAs. The Human Immunodeficiency Virus type 1 (HIV-1) has a complex RNA splicing pattern. Its genome contains 5 splicing donor (SD) sites and 9 splicing acceptor (SA) sites (1, 2), that are used in various combination, generating at least 40 different singly and multiply spliced mRNA species (1-6). The multiply spliced mRNAs encode the Tat and Rev regulatory proteins and Nef. The singly spliced molecules encode proteins Env, Vif, Vpr and Vpu. The unspliced RNA is used to produce the Gag and Pol precursors and is packaged into newly synthesized virions. Hence, for an efficient production of virions, a balanced production of these various mRNAs has to be maintained. This balance depends upon : i) the peculiarities of the HIV-1 SD and SA sites, and ii) the production of the viral Rev protein, that promotes the cytoplasmic accumulation of singly spliced and unspliced mRNAs (7). As the Rev action seems to depend upon the peculiarities of the HIV-1 SA sites (8, 9), a complete understanding of the Rev mechanism of action and of the complex alternative splicing of HIV-1 RNA requires a detailed analysis of HIV-1 SA sites.

One general peculiarity of these sites, that may in part explain their low efficiency, is the short length of the polypyrimidine tracts and their interruption by purines (10). Utilization of the A3 and A7 SA sites was also found to be inhibited by *cis* acting elements, called SRE (11-13). Furthermore, an unusual utilization of a U residue at the branch site was found for site A7 (14). Whereas detailed analysis were made for sites A3 and A7, no deep analysis was made for the SA site A2, that is responsible for the production of the Vpr mRNAs (5). In particular, efficiency of site A2 was never tested *in vitro. In vivo,* site A2 is coupled with the D1 or the D2 SD site. To fulfill the gap of information, we made two constructs for the *in vitro* production of RNAs with the D1-A2 and the D2-A2 pairs of splice sites. The *in vitro* splicing efficiencies of the two transcripts were compared to that of a previously described efficiently spliced human adenovirus-2 RNA transcript (15). The branch site (BS) used at site A2 was identified and found to be unusual. The effect of the branch site sequence on *in vitro* splicing efficiency was also tested by site-directed mutagenesis.

MATERIALS AND METHODS

Plasmid construction. The source of HIV-1 cDNA sequences from the BRU/LAV isolate was plasmid pBRU3 (16). PCR amplifications were according to Nour et al. (17). Fragment 100 [HIV-1 BRU cDNA from position 4496 to 5626 according to the numbering of Ratner et al. (18)], was amplified with a sense primer O-1 (5′ TATTCT*GGATCC*GACAGCAGAGAT 3′) that hybridizes from position 4496 to

[1] Present address: Laboratoire de Biologie Cellulaire du Développement, Université Henri Poincaré Nancy 1, Boulevard des Aiguillettes, BP239, 54506 Vandoeuvre les Nancy Cedex, France.
[2] Corresponding author. Fax: 33 3 83 91 20 93. E-mail: cbranlant@legg.u-nancy.fr.

0006-291X/97 $25.00

4507 (underlined) and generates a *Bam*H I restriction site (italic), and an antisense primer O-2 (5′ TATTCT*CTGCAG*TGCTTTGATAGA 3′) that hybridizes from position 5626 to 5615 (underlined) and generates a *Pst* I restriction site (italic). To get plasmid pLD-100, the *Bam*H I-*Pst* I fragment from PCR product 100 was inserted in plasmid pBluescript KSII$^+$ (19). To generate a pre-mRNA with a mini intron D1-A2, a *Bgl* II restriction site was introduced in the middle of the D2-A2 intron by site-directed mutagenesis. Site-directed mutagenesis was made according to Kramer et al. (20) with the oligonucleotide O-3 5′ GTGGCAAGTAGATCTTGAGGATTAG 3′ (the underlined sequences are identical to the BRU RNA from positions 4652 to 4663 and 4669 to 4678, respectively), after transfer of the *Bam*H I-*Pst* I fragment in phage M13mp9. Plasmid pLD-101 was obtained by reinsertion of the mutagenized *Bam*H I-*Pst* I fragment in pBluescriptKSII$^+$.

Construct pLD-104 was obtained by deletion of the 593-nt *Sau*3A I-*Eco*R V fragment in plasmid pLD-101 (position 5049 to 5626 of the BRU cDNA plus 15 nts of the polylinker). To get plasmid pLD-104 (Fig. 1), the resulting 3.4 kb fragment of pLD-101 was blunt ended with the Klenow enzyme and circularized by the T4 DNA ligase according to Maniatis et al. (21).

To substitute the D2 SD site of construct pLD-104 by the D1 SD site, the cDNA fragment 141 (position 1 to 384 of the BRU cDNA) was PCR amplified. The sense primer O-4 (5′ TATTCT*GGATCC*GGTCTCTCTGGT 3′) had the BRU RNA sequence from position 1 to 12 (underlined) and generated a *Bam*H I restriction site (italic). The antisense primer O-5 (5′ TATTCT*AGATCT*CCCATCGATCTA 3′) was complementary to the BRU RNA from position 384 to 373 (underlined) and generated a *Bgl* II restriction site (italic). To get plasmid pLD-C2 (Fig. 1), the *Bam*H I-*Bgl* II fragment of PCR amplification product 141 was inserted in plasmid pLD-104 in place of the 171-nt *Bam*H I-*Bgl* II fragment containing site D2.

Variant plasmid pLD-C2.838 was produced by substitution of the wt BS sequence in construct pLD-C2, by a canonical BS sequence, by site-directed mutagenesis (20). The oligonucleotide O-5 (5′ GTAGACCCTGAACTACTAACCCAACTAATTCATC 3′) had a 13-nt 5′ sequence and a 14-nt 3′ sequence identical to the BRU RNA from position 4914 to 4926 and 4934 to 4947, respectively, the underlined sequence corresponds to a consensus BS sequence (22).

Plasmid pBS-837D was obtained by insertion of the 837-pb *Pst* I - *Eco*R I fragment from plasmid pSP64-837D (15), that contains the D1-A2 splice sites of the E3 transcription unit from human adenovirus-2, into pBluescriptKSII$^+$ cleaved by the same enzymes.

In vitro transcription. Prior to transcription with T7 RNA polymerase, plasmid pLD-104 was digested with the *Hin*d III nuclease, plasmids pLD-C2, pLD-C2.838 and pBS-837D with the *Eco*R I nuclease. Conditions for production of uniformally labeled capped RNA transcripts were as follows : 250 ng of linearized DNA template were incubated for 1h in 10 μl of transcription mixture containing 6 mM $MgCl_2$, 10 mM NaCl, 40 mM Tris HCl (pH 7.9), 10 mM DTT, 0.1 mg/ml BSA, 25 U of RNasin, 1.25 mM GpppG, 0.05 mM UTP, 0.5 mM ATP and CTP, 0.2 mM GTP, 0.9 μl of [α-^{32}P] UTP (800 Ci/mmol, Dupont de Nemours) and 10 units of T7 RNA polymerase. DNA was then digested with 2 U of RNase-free DNase I. RNA was phenol extracted and ethanol precipitated in the presence of 20 μg of a yeast tRNA mixture. For branch-point analysis, large quantities of cold transcripts were synthesized with the same procedure, with 30 μl of the transcription mixture described above, where [α-^{32}P] UTP was replaced by 0.5 mM UTP.

In vitro splicing and gel electrophoresis. The HeLa cell nuclear extracts were purchased from the Computer Cell Culture Center S.A. (Belgium). Each splicing reaction was carried out with 10^5 Cerenkov cpm of RNA transcript. To get a splicing kinetics, incubations were for 30min, 1h, 1h30 and 2h30 at 30°C, in 22 μl of splicing solution containing : 53 mM KCl, 2.8 mM $MgCl_2$, 20 μM EDTA, 1 mM ATP, 28 mM creatine phosphate, 0.3 mM DTT, 3.5% (w/v) polyvinyl alcohol, 10.5% (v/v) glycerol, 79 mM Hepes (pH 7.9), 10 U of RNasin and 8 μl of nuclear extract. The spliced products were deproteinized by a 30 min treatment with proteinase K (23), followed by phenol and chloroform extractions and they were analyzed by electrophoresis on a 5% polyacrylamide sequencing gel (21). Splicing efficiency was estimated by scanning the gel with a Molecular Dynamics Phosphorimager, using the Image Quant software version 3.3. The M/P ratio (amount of mature RNA versus the amount of residual precursor) was determined taking into account the measured radioactivity and the uridine content of the molecules.

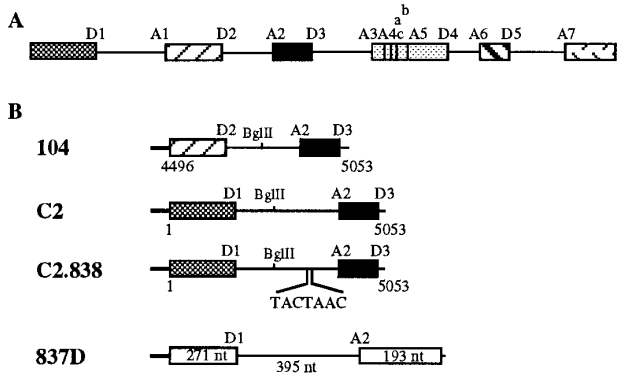

FIG. 1. Schematic representation of the RNA transcripts used in this study. Exon-intron organization of the HIV-1 RNA (5) (A) and RNA transcripts 104, C2, C2.838, and 837D (B). Exons are represented by rectangles of different colors, introns by thin lines. Nucleotide positions in the HIV-1 BRU RNA (19) are indicated. The thick lines correspond to the 52-nt sequence generated by the vector. The mutated BS sequence is shown for transcript C2.838.

Branch site identification. For BS identification, the cold RNA transcript synthesized with 30 μl of transcription mixture was divided in two portions : two third were subjected to a splicing reaction for 50 min at 30°C, with 72 μl of nuclear extract in 198 μl of the splicing solution described above. To get a control, the remaining third was incubated with 24 μl of nuclear extract in 66 μl of splicing solution, except that the ATP concentration used was of 8 mM instead of 1 mM, and that the incubation temperature was of 33°C instead of 30°C.

After incubation, the control RNA and the spliced RNA were treated by proteinase K, phenol extracted and ethanol precipitated. One half of the spliced RNA was subjected to the debranching activity of 10 μl of nuclear extract, in a 60 μl solution containing : 6.5 mM EDTA, 66 mM KCl, 0.33 mM DTT, 13.3% glycerol, 13.3 mM Hepes pH 8.0, for 30 min at 30°C. RNA was again extracted with a phenol-chloroform mixture and ethanol precipitated. Position of the branch nucleotide was analyzed by primer extension with reverse transcriptase (RT), using primer O-6 5′ CACCTAGGGCTAACTATATG 3′, complementary to the BRU RNA from position 4990 to 5009. Conditions for primer extension analysis were as previously described (24). The sequencing ladder was prepared according to Qu et al. (25), with the 5′-end labeled oligonucleotide O-6. The cDNA products were fractionated by electrophoresis on a 7% sequencing gel (21).

RESULTS

In Vitro Splicing Efficiency of the D1-A2 and D2-A2 Splice-Site Pairs

As the HIV-1 A2 SA site is used with the D1 or D2 SD sites *in vivo*, two DNA constructs were produced to test the *in vitro* efficiency of site A2 in a HIV-1 sequence context. Construct pLD-104 (Fig. 1) was made to test

the efficiency of the D2-A2 pair. It codes for a mRNA with the HIV-1 BRU sequence from position 4496 to 5053, with mutations in the intron that create a *Bgl* II restriction site. The 613-nt transcript 104 produced by *in vitro* trancription of construct 104 contained : i) the HIV-1 exon 2 (A1-D2) flanked by 52 additional nts, that were generated by the vector, ii) the mutated D2-A2 intron, and iii) exon 3 (A2-D3) flanked by 8 nts of the D3-A3 intron. As *in vitro* splicing assays are only efficient for pre-mRNAs with lengths that do not exceed 1600 nts, the splicing efficiency of the D1-A2 intron (4682-nt long) was tested on a truncated version of this intron coded by construct C2. The C2 RNA transcript (Fig. 1) contained : i) the exon 1 (position 1 to 289), flanked by 52 nts generated by the vector, ii) the 96-nt sequence located downstream from site D1 joined to the 308-nt sequence located upstream of site A2 and iii) the entire exon 3 flanked by 8 nts of the D3-A3 intron.

Uniformely labeled transcripts 104 and C2 were incubated for 30 min, 1h, 1h30 and 2h30 in a HeLa cell nuclear extract in conditions previously defined to get an efficient *in vitro* splicing of the human adenovirus-2 control RNA transcript (15). Fractionation of the resulting products by electrophoresis on a 5% sequencing gel, revealed an efficient splicing of transcripts 104 and C2, as compared to the control RNA (Fig. 2). For transcript 104, both the 187-nt mature RNA (band 5), the intron in a lariat form (band 2) were accumulated during the incubation. Band 3 probably corresponds to a degradation product of the intron retaining the lariat structure and band 1 to the first intermediate containing the lariat. The first exon was lost during the electrophoresis. For transcript C2, a very similar pattern was observed : the first intermediate (band 1), the intron in a lariat form (band 3), a degradation product of the intron retaining the lariat form (band 4) and the 425-nt mature RNA (band 5) were also accumulated after a 2h30 incubation, the free exon 1 (band 6) was retained in the gel in this case. Based on the evaluation with a Phosphorimager of the radioactivity contained in the bands of gel corresponding to the residual precursor RNA and the mature RNA obtained after a 2h30 incubation, M/P ratios of 0.7 and 1.6 were found for transcripts 104 and C2, respectively. The experience was repeated several times and M/P ratios comprised between 0.7 and 1 and between 1.4 and 2 were reproducibly found for transcripts 104 and C2, respectively. An M/P ratio of 1 was found for the control adenovirus transcript. Hence, the HIV-1 SA site A2 is efficiently used *in vitro* and the efficiency of the D1-A2 pair is superior to that of the control adenovirus splice-site pair.

Utilization of an Unusual Branch Point at Site A2

To check whether an unusual residue was used at site A2, as found for site A7 (14), the cold C2 transcript was incubated in splicing conditions for 50 min. This short incubation time was chosen to optimize the yield of branched molecules (Fig. 2). The BS was identified by direct primer-extension analysis of the splicing products. For identification of specific arrest of RT at the branched residues, on the one hand, half of the spliced products was subjected to the debranching activity of a nuclear extract. On the other hand, C2 transcript was incubated in splicing conditions, except that in order to block the splicing reaction, the ATP concentration was of 8 mM instead of 1 mM and the incubation temperature of 33°C instead of 30°C. The fractionation of the cDNAs obtained in the control experiment and in the assays is illustrated in Figure 3. Two successive RT arrests were detected with the splicing products. They were absent in the control RNA and their yields decreased slightly after the debranching step.

The arrest located 38 nts upstream of site A2 was of medium intensity, that located one nucleotide upstream was very strong. As RT is expected to stop predominantly one nucleotide before the branched nucleotide (26), the results strongly suggest that the G residue located 40 nts upstream of site A2 is used as the main BS. Either some of the RT molecules stopped two nts before the branched G residue, explaining the second arrest, or the A residue that follows it is also used as a BS. Identical results were obtained when site A2 was coupled with site D1 (not shown). Hence, the A2 BS sequence is UAGCAGA with a preferential utilization of the underlined G residue as the branched nucleotide.

Effect of the BS Sequence on in Vitro Splicing Efficiency

To test whether the conversion of the UAGCAGA BS sequence of site A2 into a canonical UACUAAC BS sequence will increase the utilization of this site, a mutated construct, that produced a C2.838 transcript with a canonical BS sequence located at the same position as the UAGCAGA wt sequence, was generated by site-directed mutagenesis. As shown in Figure 3A, the splicing efficiency of transcript C2.838 was not significantly increased as compared to that of transcript C2 (M/P ratio of 2.1 for C2.838 against 1.6 for C2). Upon primer extension analysis of the splicing products from transcript C2.838 with or without debranching step, two RT stops were detected at the same positions as for the wt C2 transcript (38 and 39 nts upstream of site A2). The position of the strongest stop is in accord with the expected utilization of the ultimate A residue of the canonical BS sequence (UACUAAC) as the branched nucleotide (Fig. 3B). The more likely explanation for the second band of lower intensity is the arrest of some of the RT molecules two nucleotides before the branched nucleotide.

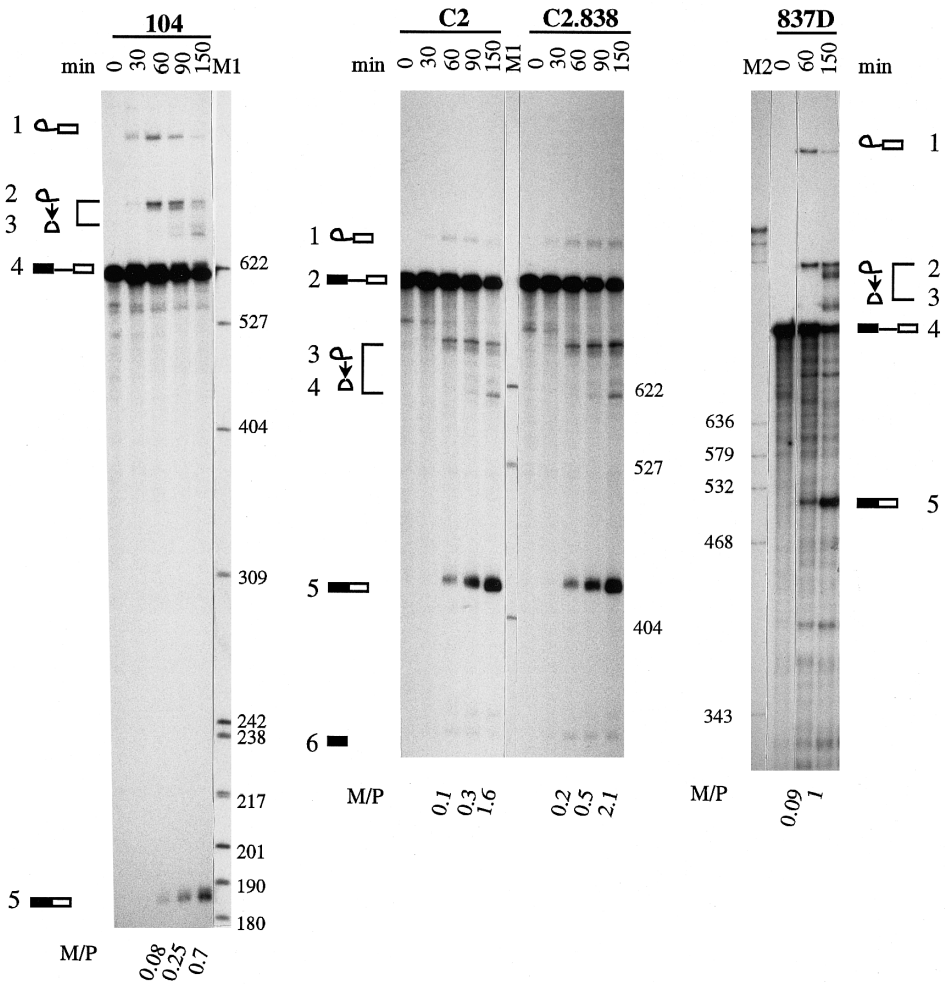

FIG. 2. *In vitro* splicing kinetics of transcripts 104, C2, C2.837 and 837D. Uniformely labeled transcripts 104, C2 and C2.838 were incubated for 0, 30 min, 1h, 1h30 and 2h30 in a HeLa cell nuclear extract as described in Materials and Methods. The control transcript 837D was spliced for 1h and 2h30 in the same conditions. The splicing products were fractionated on a 5% sequencing gel and autoradiographed. Name of transcripts and duration of incubations are indicated above the lanes and M/P values evaluated as described in Materials and Methods at the bottom of the lanes. Positions of the precursor RNA and splicing products are shown on the left. Labeled size markers M1 and M2 were pBR322/*Hpa* II and λ/*Pvu* II, respectively. Fragment lengths are indicated in nt.

DISCUSSION

Removal of introns from precursors to messenger RNAs involves two trans-esterification reactions and takes place in a large macrostructure, the spliceosome (for review 27). Except for a very limited number of introns, in the first trans-esterification, the 2′ OH of an adenosine residue is the nucleophile that attacks the 5′-splice site phosphate (reviewed in 27). Here we show that a guanosine residue is preferentially used as the nucleophile at the HIV-1 SA site A2. Upon varying the identity of the base in a canonical BS sequence, Query et al. (28, 29) found that the highest efficiency of A residue at the BS, as compared to G or U residues, is due in part to the strong inhibitory effect of the C6-oxo/ N1-H of guanine and uracile. In addition, the hydrogen bond acceptor properties of the N1 position of adenine could be involved in the second trans-esterification step (29). In spite of this more favorable properties of A residue to assume the function of branch nucleotide, we observed, here, that a D1-A2 intron with a canonical BS sequence is not significantly more efficiently spliced *in vitro,* than a wt D1-A2 intron with a G residue at the BS. One possibility is that the first trans-esterification step is not the limiting step in the splicing reaction of the D1-A2 intron. However, it should be noticed that, as compared to the adenovirus RNA transcript used as the control, the C2 and 104 transcripts were efficiently spliced *in vitro* (Fig. 2). Hence, introns D1-A2 or D2-A2 with a suboptimal polypyrimidine tract and an unusual BS sequence are, however, efficiently assembled into spliceosomes and excised. Interestingly, the sequence that we identified as the A2 BS sequence in the BRU isolate is highly

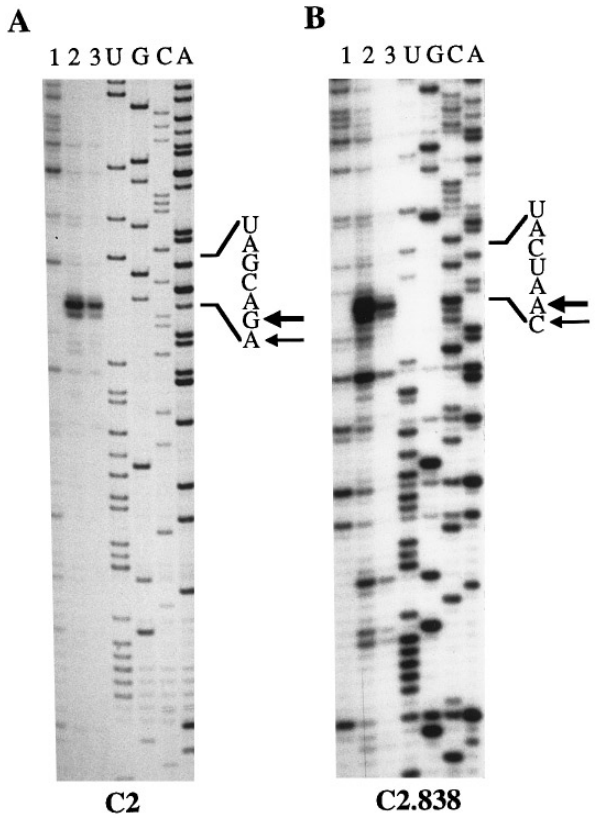

FIG. 3. Branch point identification for transcript C2 and C2.838. Cold transcripts C2 (A) and C2.838 (B) were subjected to a 50 min splicing reaction as described in Materials and Methods. One half was analyzed directly by primer extension with oligonucleotide O-6 (Lane 2), the other half was subjected to the debranching activity of a nuclear extract (Lane 3) prior primer extension analysis. Lane 1 corresponds to a control primer extension on transcripts incubated for 50 min at 33°C in a nuclear extract containing 8 mM ATP. Lanes UGCA correspond to the sequencing ladder. The BS sequence and the RT arrests are shown.

conserved in various HIV-1 isolates. So that, the peculiarity that we observed for BRU seems to be general. Except for a few cases, the experimentally identified BS were found to be located between the 18th and the 40th nt upstream of the 3′-splice site (30-32). Hence, concerning its distance to the 3′-splice site (39 nts), the BS identified for site A2 fits to the general rule.

In addition to its distance from the 3′-splice site, the identity of the nucleotide used as the BS is also defined by the capacity of the BS sequence to base pair with U2 snRNA (33-35). The branched nucleotide constitutes a bulged residue within the interaction (29). In contrast to the situation in yeast (36, 37), the BS sequence is only weakly conserved in metazoans and as a consequence the base-pair interaction with U2 snRNA is often of low stability (Fig. 4). However, the canonical UACUAAC sequence was found to be preferentially used (22). A rather stable interaction can be formed between the A2 BS sequence and U2 snRNA.(Fig. 4). Assuming that the two adenosines are stacked together, the branched G residue is a bulged nucleotide. The interaction involves 5 Watson- Crick base pairs against 6 for a canonical BS sequence. This may in part explain the observed *in vitro* splicing efficiency. Dyhr-Mikkelsen and Kjems (14) also found the utilization of an unusual U residue as the BS for the HIV-1 SA site A7. As shown in figure 4, a base-pair interaction with U2 snRNA with a stability similar to that of the interaction formed at the A2 BS, can be formed at the A7 BS. A very limited number of cellular pre-messenger introns using residues other than adenosine as BS have been described. A C residue was found to be used for the intron A of the human growth hormone (38). A uridine branch residue was also found to be a cis-acting element involved in regulation of the alternative processing of the calcitonin/CGRP-I pre-mRNA (39) and substitution of the uridine residue at the BS by an adenosine resulted in a strong increase of the calcitonin specific splicing efficiency (39). The peculiar BS sequence of the HIV-1 SA sites may also be a basis for the regulation of their relative utilization. Experiments are underway to identify the BS sequences of the A1 and A3 SA sites.

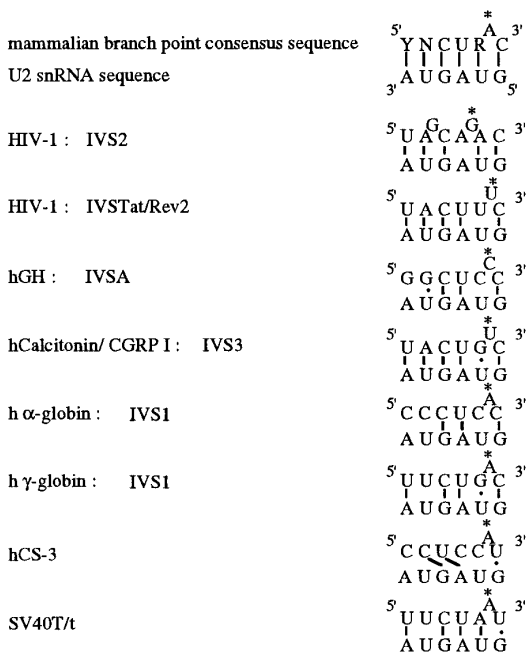

FIG. 4. Possible base-pair interaction between U2 snRNA and some of the BS sequences identified for metazoans. The interaction between the metazoan consensus BS sequence and U2 snRNA is shown, as well as, the possible interaction for the two BS sequences identified in HIV-1 RNA (site A2 this paper), site A7 (15), the BS sequence of the intron A of the human growth hormone (38), the intron 3 of the calcitonin/CGRP I (39), and some of the identified BS sequences with a branched A residue (40). The branched residues are indicated by an asterisk.

ACKNOWLEDGMENTS

F. Clavel is thanked for providing plasmid pBRU3. L. Damier was a fellow of the French Ministère de la Recherche et de l'Enseignement Supérieur and then from the SIDACTION foundation. This work was supported by the Association Nationale pour la Recherche contre le SIDA (ANRS), the Centre National de la Recherche Scientifique, and the French Ministère de la Recherche et de l'Enseignement Supérieur.

REFERENCES

1. Furtado, M. R., Balachandran, R., Gupta, P., and Wolinsky, S. M. (1991) *Virology* **185,** 258–270.
2. Schwartz, S., Felber, B. K., Benko, D. M., Fenyo, E. M., and Pavlakis, G. N. (1990) *J. Virol.* **64,** 2519–2529.
3. Purcell, D. F., and Martin, M. A. (1993) *J. Virol.* **67,** 6365–6378.
4. Robert-Guroff, M., Popovic, M., Gartner, S., Markham, P., Gallo, R. C., and Reitz, M. S. (1990) *J. Virol.* **64,** 3391–3398.
5. Schwartz, S., Felber, B. K., and Pavlakis, G. N. (1991) *Virology* **183,** 677–686.
6. Schwartz, S., Felber, B. K., Fenyo, E. M., and Pavlakis, G. N. (1990) *J. Virol.* **64,** 5448–5456.
7. Sodroski, J., Goh, W. C., Rosen, C., Dayton, A., Terwilliger, E., and Haseltine, W. (1986) *Nature* **321,** 412–417.
8. Chang, D. D., and Sharp, P. A. (1989) *Cell* **59,** 789–795.
9. Mikaelian, I., Krieg, M., Gait, M. J., and Karn, J. (1996) *J. Mol. Biol.* **257,** 246–264.
10. Staffa, A., and Cochrane, A. (1994) *J. Virol.* **68,** 3071–3079.
11. Amendt, B. A., Si, Z. H., and Stoltzfus, C. M. (1995) *Mol. Cell. Biol.* **15,** 4606–4615.
12. Amendt, B. A., Hesslein, D., Chang, L. J., and Stoltzfus, C. M. (1994) *Mol. Cell. Biol.* **14,** 3960–3970.
13. Staffa, A., and Cochrane, A. (1995) *Mol. Cell. Biol.* **15,** 4597–4605.
14. Dyhr-Mikkelsen, H., and Kjems, J. (1995) *J. Biol. Chem.* **270,** 24060–24066.
15. Domenjoud, L., Gallinaro, H., Kister, L., Meyer, S., and Jacob, M. (1991) *Mol. Cell. Biol.* **11,** 4581–4590.
16. Charneau, P., Alizon, M., and Clavel, F. (1992) *J. Virol.* **66,** 2814–2820.
17. Nour, M., Naimi, A., Beck, G., and Branlant, C. (1995) *Curr. Microbiol.* **31,** 270–278.
18. Ratner, L., Haseltine, W., Patarca, R., Livak, K. J., Starcich, B., Josephs, S. F., Doran, E. R., Rafalski, J. A., Whitehorn, E. A., Baumeister, K., Ivanoff, L., Petteway, S. R., Pearson, M. L., Lautenberger, J. A., Papas, T. S., Ghrayeb, J., Chang, N. T., Gallo, R. C., and Wong-Staal, F. (1985) *Nature* **313,** 277–283.
19. Short, J. M., Fernandez, J. M., Sorge, J. A., and Huse, W. D. (1988) *Nucleic Acids Res.* **16,** 7583–7600.
20. Kramer, B., Kramer, W., and Fritz, H. J. (1984) *Cell* **38,** 879–887.
21. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
22. Zhuang, Y. A., Goldstein, A. M., and Weiner, A. M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 2752–2756.
23. Krainer, A. R., Maniatis, T., Ruskin, B., and Green, M. R. (1984) *Cell* **36,** 993–1005.
24. Mougin, A., Gregoire, A., Banroques, J., Segault, V., Fournier, R., Brule, F., Chevriermiller, M., and Branlant, C. (1996) *RNA - A Publication of the RNA Society* **2,** 1079–1093.
25. Qu, H. L., Michot, B., and Bachellerie, J. P. (1983) *Nucleic Acids Res.* **11,** 5903–5920.
26. Ruskin, B., Greene, J. M., and Green, M. R. (1985) *Cell* **41,** 833–844.
27. Moore, M. J., Query, C. C., and Sharp, P. A. (1993) *in* The RNA World (Gesteland, R., and Athuis, J., Eds.), pp. 303–358, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
28. Query, C. C., Moore, M. J., and Sharp, P. A. (1994) *Genes Dev.* **8,** 587–597.
29. Query, C. C., Strobel, S. A., and Sharp, P. A. (1996) *EMBO J.* **15,** 1392–1402.
30. Ruskin, B., Krainer, A. R., Maniatis, T., and Green, M. R. (1984) *Cell* **38,** 317–331.
31. Reed, R., and Maniatis, T. (1985) *Cell* **41,** 95–105.
32. 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.
33. Parker, R., Siliciano, P. G., and Guthrie, C. (1987) *Cell* **49,** 229–239.
34. Wu, J., and Manley, J. L. (1989) *Genes Dev.* **3,** 1553–1561.
35. Zhuang, Y., and Weiner, A. M. (1989) *Genes Dev.* **3,** 1545–1552.
36. Langford, C. J., Klinz, F. J. C., Donath, D., and Gallwitz, D. (1984) *Cell* **36,** 645–653.
37. Rodriguez, J. R., Pikielny, C. W., and Rosbash, M. (1984) *Cell* **39,** 603–610.
38. Hartmuth, K., and Barta, A. (1988) *Mol. Cell. Biol.* **8,** 2011–2020.
39. Adema, G. J., van Hulst, K. L., and Baas, P. D. (1990) *Nucleic Acids Res.* **18,** 5365–5373.
40. Nelson, K. K., and Green, M. R. (1989) *Genes Dev.* **3,** 1562–1571.