MOLECULAR AND CELLULAR BIOLOGY, Sept. 1991, p. 4581–4590
0270-7306/91/094581-10$02.00/0

Vol. 11, No. 9

Identification of a Specific Exon Sequence That Is a Major Determinant in the Selection between a Natural and a Cryptic 5′ Splice Site

LIONEL DOMENJOUD, HÉLÈNE GALLINARO, LILIANE KISTER, SYLVIE MEYER,† AND MONIQUE JACOB*

Laboratoire de Génétique Moléculaire des Eucaryotes du Centre National de la Recherche Scientifique, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'Institut National de la Santé et de la Recherche Médicale, Institut de Chimie Biologique, Faculté de Médecine, 11, Rue Humann, 67085 Strasbourg Cédex, France

Received 4 March 1991/Accepted 24 June 1991

The first intron of the early region 3 from adenovirus type 2 contains a cryptic 5′ splice site, Dcr1, 74 nucleotides downstream from the natural site D1. The cryptic site can be activated when the natural site is inactivated by mutagenesis. To investigate the basis for selection between a natural and a cryptic 5′ splice site, we searched for *cis*-acting elements responsible for the exclusive selection of the natural site. We show that both the relative intrinsic strength of the sites and the sequence context affect the selection. A 120-nucleotide segment located at the 3′ end of exon 1 enhances splicing at the proximal site D1; in its absence the two sites are used according to their strength. Thus, three *cis*-acting elements are involved in the silencing of the cryptic site: the sequence of D1, the sequence of Dcr1, and an upstream exonic sequence. We show that the exonic element folds, in solution, into a 113-nucleotide-long stem-loop structure. We propose that this potential stem-loop structure which is located 6 nucleotides upstream of the exon 1-intron junction is responsible for the preferential use of the natural 5′ splice site.

Regulation of primary transcript processing is one of the key events ensuring the production of functional mRNA. Processing mainly includes splicing and polyadenylation (for reviews, see references 5, 30, and 50). In some genes, alternative splicing, polyadenylation, or both occur, which may produce a variety of mRNAs with subtle different characteristics. Usually, the various RNA forms are either made or predominate in different cell types or at different stages of the cell's life.

We are interested in the alternative processing events occurring in the region of adenovirus type 2 (Ad2) where the transcription units of early region 3 (E3) and late region 4 (L4) overlap (18 and Fig. 1). Processing implies the selection of either the early splicing sites or the late polyadenylation site. In the course of our in vitro studies, mutagenesis experiments revealed the presence of a cryptic 5′ splice site (Dcr1) located downstream of the natural 5′ splice site D1 of the first E3 intervening sequence (IVS1).

In mammals, the 5′ splice site (or donor site) consensus sequence is $^{C}_{A}$AG/GU$^{A}_{G}$AGU, including three exonic and six intronic bases (4, 36), and is recognized by a conserved sequence at the 5′ end of U1RNA. In fact, this consensus masks a large diversity of individual sequences (26) and only the first nucleotide (G1) is conserved in all natural introns. All other nucleotides are dispensable for function provided that the overall stability of the U1RNA-donor site hybrid is conserved (26, 60). From the donor site sequence diversity, it can be predicted that potential donor sites will be present fairly often along an RNA strand. GU is one of the 16 dinucleotides; thus, statistically, there are 62.5 GUs per 1,000 nucleotides. Some of these may be eliminated as potential donor sites because they cannot pair stably with U1RNA, and others may be eliminated because there is no suitable acceptor or branch site. Nevertheless, some of these sequences are donor site candidates, in particular when they are located close to the natural donor site and can use the same acceptor and branch sites. The best-known example is that of the human β-globin gene where the donor site of the first IVS is surrounded within a 60-nucleotide (nt) stretch by three cryptic sites used only when the natural site is mutated (56). Thus, most of the natural pre-mRNAs may contain cryptic donor sites in addition to constitutive or alternative sites. This is the case for the adenovirus region that we are studying, where the cryptic site Dcr1, located 74 nt downstream of the natural site D1, can use the natural acceptor and branch sites. In order to understand the basis for splice site selection, we used the E3 pre-mRNA of Ad2 as a model system for the study of the mechanism of selection between a natural and a cryptic donor site.

How does the cell achieve selection between the natural and cryptic sites? It has been shown that the intrinsic strength of the site, i.e., its capacity to form a stable hybrid with U1RNA, is a major element for selection (14, 39). However, two lines of evidence suggest that intrinsic strength is not the only element of selection between donor sites. First, if it were, each site would be used according to its strength, which is obviously not so in the case of cryptic sites. Second, the importance of sequence context, in particular of exon sequences, for selection of natural sites in alternative splicing has been shown in a number of systems (10, 11, 17, 21, 27, 31, 33, 38, 43, 45, 47, 52, 54, 57) and this might also apply for selection between a natural and a cryptic site. Therefore, we looked to see whether *cis*-acting elements other than the splice sites themselves were involved in selection. We found an upstream exonic sequence whose deletion provoked the activation of the cryptic site even in the presence of the natural donor site. We also showed that, in the absence of the exonic sequence, the two sites were used according to their relative strength but that, in its presence, splicing at the natural site was enhanced.

* Corresponding author.
† Present address: F. Hoffmann-La Roche & Co., Aktiengesellschaft, CH.4002 Basel, Switzerland.

Finally, we demonstrate that the upstream exonic sequence with enhancing activity can be folded into a stem-loop structure.

MATERIALS AND METHODS

Plasmid construction. The Ad2 *Eco*RI-D DNA fragment (23) containing the 5′ part of the E3 unit was digested with *Pst*I and *Sac*I, and the *Pst*I-*Sac*I DNA fragment (837 bp) was inserted in the polylinker of pSP64 (35). The plasmid is referred to as pSPWT. For the construction of 5′ deletion mutants, the pSPWT plasmid was linearized at the *Pst*I site and treated with exonuclease *Bal* 31 (Bethesda Research Laboratories) in order to remove from 50 to 300 bp from the 5′ end of the E3 sequence. After the ends were filled in with the Klenow enzyme, the E3 residual sequences were excised by *Eco*RI digestion and purified by electrophoresis in 1% agarose gels. The E3 DNA fragments were ligated by T4 DNA ligase (Appligène) to an pSP64 vector opened by *Hin*cII and *Eco*RI. Mutants 5′dl 41, 5′dl 59, 5′dl 62, 5′dl 37, 5′dl 36, and 5′dl 67 (see Fig. 3) were prepared in this way. For the construction of 5′dl 94, a *Pst*I-*Sac*II fragment was removed from pSPWT. After filling in with the Klenow enzyme, the plasmid was recircularized by T4 DNA ligase.

Oligonucleotide-directed mutagenesis was performed on double-stranded plasmid DNA by the method of Inouye and Inouye (25). Mutants pm1 to pm6 were obtained by this method.

The extent of deletions and the point mutations were determined by using the dideoxynucleotide sequencing procedure (62).

RNA synthesis. The plasmids used as DNA templates were first linearized by *Eco*RI. In vitro transcription and capping was performed by the method of Krainer et al. (28) in the presence of [α-^{32}P]CTP (~400 Ci/mmol). After treatment with RNase-free DNase I and deproteinization (19), the transcripts were ethanol precipitated in the presence of yeast tRNA as a carrier. The integrity of transcripts was checked by electrophoresis in 5% polyacrylamide–8 M urea. The transcripts harbor one polylinker nucleotide at their 3′ ends and a variable number of plasmid nucleotides at their 5′ ends, depending on the cloning procedure.

In vitro splicing and RNA analysis. Nuclear extracts were prepared from S3 HeLa cells by the method of Dignam et al. (12). The in vitro splicing reactions were carried out essentially as described previously (28). They were performed in a total volume of 25 μl including 10 μl of nuclear extract and ~15,000 Cerenkov cpm of ^{32}P-RNA transcripts. The final concentration of the other components was as follows: 13 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-ethanesulfonic acid) (pH 7.9), 60 mM KCl, 2 mM $MgCl_2$, 20 μM EDTA, 1 mM ATP, 20 mM creatine phosphate, 0.5 mM dithiothreitol, 2.6% (wt/vol) polyvinyl alcohol, 14% (vol/vol) glycerol, and 400 U of RNasin per ml. Standard incubation was for 3 h at 30°C. The processed RNA was purified by proteinase K treatment and phenol-chloroform extraction (28), and the RNA was analyzed in 5 or 6% polyacrylamide gels (acrylamide/bisacrylamide, 20/1 [wt/wt]) containing 8 M urea. After drying, the gels were autoradiographed with an intensifying screen at −70°C. The autoradiographs were scanned (Vernon recorder, Paris, France) or alternatively, the bands corresponding to the different molecules were cut out and their radioactivity was counted. Splicing efficiency was $b/(b + a)$, b being the sum of splicing intermediates and products and a being the residual precursor transcript.

Secondary structure of transcripts in solution. Secondary structure determination involved a statistical cleavage of the RNA (1 cut per molecule) by enzymes specific for double-stranded (V1) or single-stranded (T1, micrococcal nuclease) regions, followed by identification of the cleavage products by primer extension analysis (13).

About 40,000 Cerenkov cpm of RNA was solubilized in 25 μl of splicing medium. The medium was supplemented by 1 mM $CaCl_2$ (final concentration) in the case of micrococcal nuclease digestion. The enzymes were then titrated to determine the concentration allowing statistical cleavage. Routinely, at least two enzyme concentrations (around the optimal concentration) were tested. Digestion was for 15 min at 25°C in the presence of either 1×10^{-2} to 5×10^{-2} U of T1 RNase, 1×10^{-4} to 5×10^{-4} U of micrococcal nuclease (Worthington Biochemical Corp.), or 0.225×10^{-2} to 0.45×10^{-2} U of V1 nuclease (Pharmacia). Digestion was stopped by the addition of 25 μl of 200 mM Tris-HCl (pH 8.3)–30 mM EDTA–2% (wt/vol) sodium dodecyl sulfate–5 μg of yeast tRNA. RNA was phenol-chloroform extracted and ethanol precipitated. Control samples were not digested by nucleases but were treated identically throughout.

A primer, complementary to the sequence between positions 390 and 417 of the WT transcript, was 5′ end labeled with [α-^{32}P]ATP and T4 polynucleotide kinase. About 50,000 Cerenkov cpm was coprecipitated with each RNA sample to be analyzed. Hybridization and extension in the presence of avian myeloblastosis virus reverse transcriptase (Pharmacia) were performed by the method of Schmitt et al. (46). The synthesis of cDNA was initiated at 37°C for 5 min with 4 U of enzyme per sample and continued for 40 min at 41°C. At the end of the reaction, the template RNA was hydrolyzed by incubation in the presence of 50 μl of 0.3 M NaOH at 37°C for 1 h. After neutralization, the cDNA products were ethanol precipitated and electrophoresed in 5% polyacrylamide–8 M urea sequencing gels. The control sequence was determined from the double-stranded plasmid sequence. The same 5′-end-labeled primer described above was used and extended in the presence of dideoxynucleotides (8).

RESULTS

Description of the system. Figure 1a is a schematic representation of the region of the Ad2 genome that we are studying. The main landmarks of the overlapping E3 and L4 transcription units are indicated. The E3 IVS1, from D1 to A2, is excised for the formation of early E3 mRNAs, whereas the late polyadenylation site, located in the central part of E3 IVS1, is used for the 5′ end formation of late pVIII mRNA (53).

For in vitro studies, we constructed a synthetic transcript (Fig. 1b) containing the natural E3 IVS1 (395 nt), 256 nt of natural exon 1 (plus 15 plasmidic nucleotides), and 192 nt of natural exon 2 (plus 1 plasmidic nucleotide). The transcript is a substrate for splicing or polyadenylation when incubated under appropriate conditions, showing that the *cis*-acting elements necessary for the two processing reactions are present in the wild type (WT) transcript. Figure 2a shows a time course of the splicing reaction. After a lag period of about 20 min, the intermediates (exon 1 and lariat IVS1-exon 2) and products (mRNA and lariat IVS1) of splicing are formed. A trimming of the tail of free lariat IVS1 is observed after 60 min of incubation which is independent of the nuclear extract preparation used (six different preparations). All the molecules were characterized by various methods as required (S1 nuclease mapping, primer extension, and debranching; data not shown) and as applied previously for the

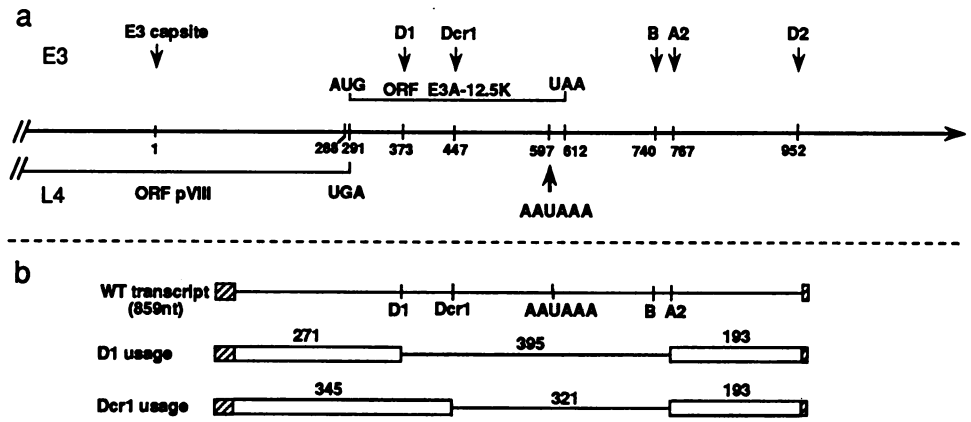

FIG. 1. (a) Structure of the region of the adenovirus genome where the transcription units E3 and L4 overlap. The landmarks (53) are indicated above the gene for E3 and below for L4. D, donor site; B, branch site; A, acceptor site. The protein encoded by open reading frame E3A-12.5K was never isolated (55), and the use of this open reading frame at the early or late period remains dubious. (b) Structure of the synthetic transcript (WT). The exon-intron structure is indicated in the case of D1 or Dcr1 usage. The hatched boxes represent plasmid fragments (15 nt at the 5′ end and 1 nt at the 3′ end). The numbers above the structure indicate the size (in nucleotides) of introns and exons, plasmidic nucleotides included.

characterization of in vivo transcripts (18, 48, 49). For our splicing studies, incubation was for 3 h. No apparent degradation of products (except for the trimming of the lariat) and no further accumulation of products was observed after this time. The optimum Mg^{2+} concentration was 2 mM. Under these conditions, splicing efficiency was 65% ± 9%.

The replacement of the first IVS nucleotide (G1) by an A which inactivates D1 activates a single cryptic donor site, Dcr1 (Fig. 1b), located 74 nt downstream (pm2, Fig. 2b). Thus, Dcr1 can function as an authentic donor site and the factors from a nuclear extract of HeLa cells can use Dcr1 but are restrained from doing so when D1 is present.

The data also suggest that there is a *cis* competition between the two sites, the competition being entirely in favor of D1 in the WT transcript. This is probably because the cryptic site is not a potent *cis*-competing element for D1 in its natural context. Indeed, the substitution of the first cryptic IVS nucleotide (G447) by an A which inactivates Dcr1 does not significantly affect D1 usage (pm1, Fig. 2b); in a set of 5 experiments, splicing efficiency at D1 was 68% ± 13% for pm1 versus 64 ± 7% for the WT transcript.

As the donor site sequence is CcG/GUGAGU for D1 and CAG/GUAgag for Dcr1 (small letters represent nucleotides which do not pair with U1RNA), the donor site-U1RNA duplexes have similar stabilities (see Fig. 4a). This argues in favor of a parameter other than intrinsic strength in the silencing of Dcr1 and in the exclusive selection of D1.

Involvement of an upstream exonic sequence in donor site selection and splicing efficiency. To search for a *cis*-acting element that may prevent Dcr1 usage, we constructed *Bal* 31 5′ deletion mutants (Fig. 3c) and tested them for splicing in vitro (Fig. 3a and b). Splicing at D1 occurs in the WT transcript, as expected, and in mutant transcripts 5′dl 41, 5′dl 59, 5′dl 62, and 5′dl 94. However, additional bands due to the activation of Dcr1 systematically appear in 5′dl 59, 5′dl 62, 5′dl 94, and occasionally 5′dl 41 (as best shown by the examination of free lariat IVS and crIVS in Fig. 3a and b). In the last three transcripts (5′dl 37, 5′dl 36, and 5′dl 67), D1 has been deleted and Dcr1 is the only donor site used. Together, the results indicate that a *cis*-acting element preventing Dcr1 usage is lacking in 5′dl 59, 5′dl 62, and 5′dl

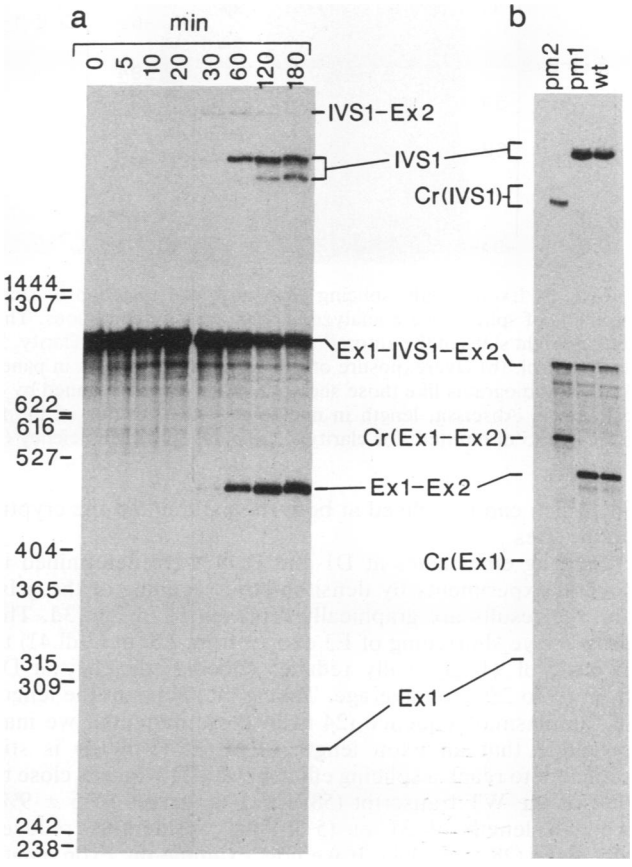

FIG. 2. (a) Time course of splicing in vitro. The WT transcript (Fig. 1b) was incubated under standard conditions. The time of incubation is indicated above the lanes. RNA was analyzed as described in Materials and Methods. The migration of size markers is indicated on the left (in nucleotide numbers). (b) Activation of Dcr1. The first natural intronic nucleotide (G373, transcript pm2) or the first cryptic intronic nucleotide (G447, transcript pm1) were replaced by an A. Incubation and analysis were under standard conditions. Identification of bands was as described previously (18, 48, 49). IVS, intact lariat intron plus products obtained by the trimming of the tail (see text); Cr, RNAs formed after cryptic site usage.

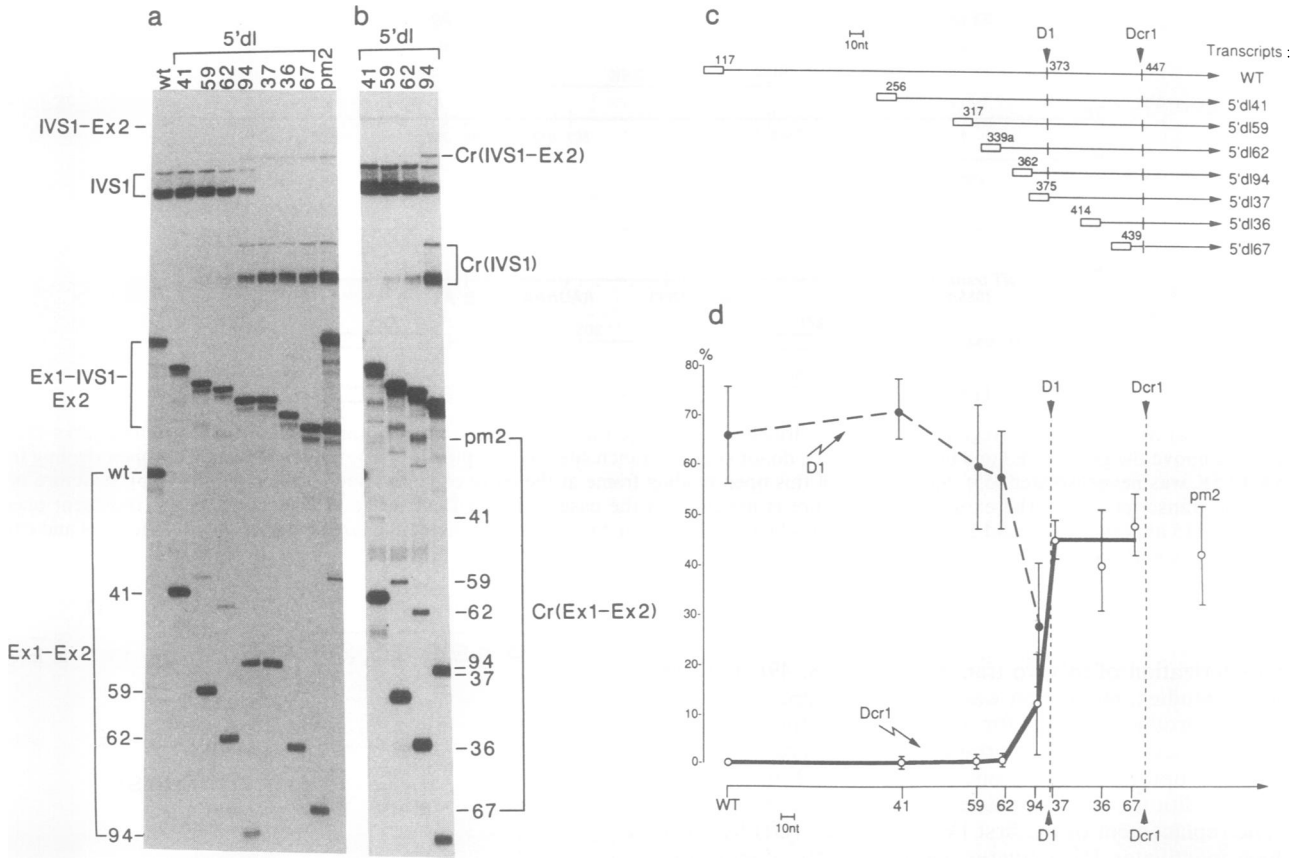

FIG. 3. Exon length, splicing efficiency, and selection of Dcr1. (a and b) The 5′ deletion mutant transcripts were incubated and the products of splicing were analyzed under standard conditions. The migration of the products of splicing at D1 and Dcr1 is indicated on the left and right sides of the autoradiogram, respectively. For clarity, 5′dl was omitted in mutant identification. Transcript pm2 (Fig. 2b) is shown as a control. (b) Overexposure of the corresponding lanes in panel a. (c) Schematic representation of the 5′ part of the deleted transcripts. (d) Autoradiograms like those shown in panel a were scanned by densitometry, and the average splicing efficiency and standard error were calculated. Abscissa, length in nucleotides; the position of the first 5′ nucleotide (plasmidic nucleotides excluded) of each transcript is indicated (5′dl omitted, for clarity). Ordinate: splicing efficiency (percentage of precursor having undergone splicing).

94, which can be spliced at both the natural and the cryptic donor sites.

Splicing efficiencies at D1 and Dcr1 were determined in several experiments by densitometric scanning of the gels, and the results are graphically represented in Fig. 3d. The progressive shortening of E3 exon 1 from 120 nt (5′dl 41) to 11 nt (5′dl 94) gradually reduces splicing efficiency at D1 from 65 to 28%, on average. Taking into account the length of the plasmid sequence (24 nt in these mutants), we may conclude that an exon length of 60 nt (5′dl 62) is still sufficient to reach a splicing efficiency at D1 which is close to that of the WT transcript (58% ± 10% versus 65% ± 9%) while a length of 31 nt (5′dl 94) considerably reduces efficiency (28% ± 13%). If we now examine the exon length required for Dcr1 usage, we observe that the cryptic exon 1 of 5′dl 67 is 31 nt long; i.e., it has the same size as the 5′dl 94 exon 1. However, in this case, splicing efficiency is not significantly different from that of 5′dl 36 or 5′dl 37 which have longer exons. Thus, exon length apparently does not have the same influence on the two donor sites. The influence of exon length on splicing efficiency was stressed previously (1, 34, 41), and the minimal exon length required for full efficiency was proposed to be 30 to 50 nt. The difference in behavior of D1 and Dcr1 strongly suggests that a feature other than length may be involved.

The progressive shortening of exon 1 also gradually activates the cryptic site Dcr1. Dcr1 splicing products were never detected with the WT transcript and only infrequently and as traces with 5′dl 41. The activation starts to be systematic in 5′dl 59 and 5′dl 62 (0.8 and 2.3%, respectively) and is always detectable upon overexposure of the gels (Fig. 3b). Finally, in 5′dl 94, efficiency of splicing at Dcr1 reached 13% on average in this set of experiments. Thus, the decrease of splicing efficiency at D1 and the activation of Dcr1 occur upon deletion of the same sequences, suggesting that the two events might be related.

Further shortening of exon 1 which eliminates the canonical GU of D1 (5′dl 37, 5′dl 36, 5′dl 67) dramatically increases the splicing efficiency at Dcr1. This in agreement with a series of observations or experiments showing that cryptic sites can be revealed upon deletion of the natural site (see Discussion). Nevertheless, splicing efficiency at Dcr1 (~44%) does not reach that obtained for D1 (~65%). It is close to that found for the mutant pm2 in which the first intronic G of D1 was replaced by an A but which conserves the exonic sequence (Fig. 3d).

In conclusion, the 3′ end of E3 exon 1 harbors a *cis*-acting element which prevents the cryptic site Dcr1 usage. The progressiveness of the changes observed (Fig. 3d) and the differences of the effect of exon length for D1 and Dcr1

a

		kcal/mol	symbol
D1	CcG |GUgAGU	-6.9	+
	CcG |aUgAGU		-
	CcG |GUgAaU	-3.1	↘
Dcr1	CAG |GUAgag	-6.2	+
	CAG |aUAgag		-
	CAG |GUAAGc	-8.8	↗

b

	D1	Dcr1	Dcr1/D1	splicing efficiencies ex1+ D1	ex1+ Dcr1	ex1- D1	ex1- Dcr1
pm1	+	-		81	-	61	-
WT	+	+	0.9	77	0	33	38
pm3	+	↗	1.3	33	6	5	66
pm4	↘	+	2.0	9	54	0	67
pm5	↘	↗	2.8	0.3	66	-	-
pm2	-	+		-	64	-	65
pm6	-	↗		-	70	-	-

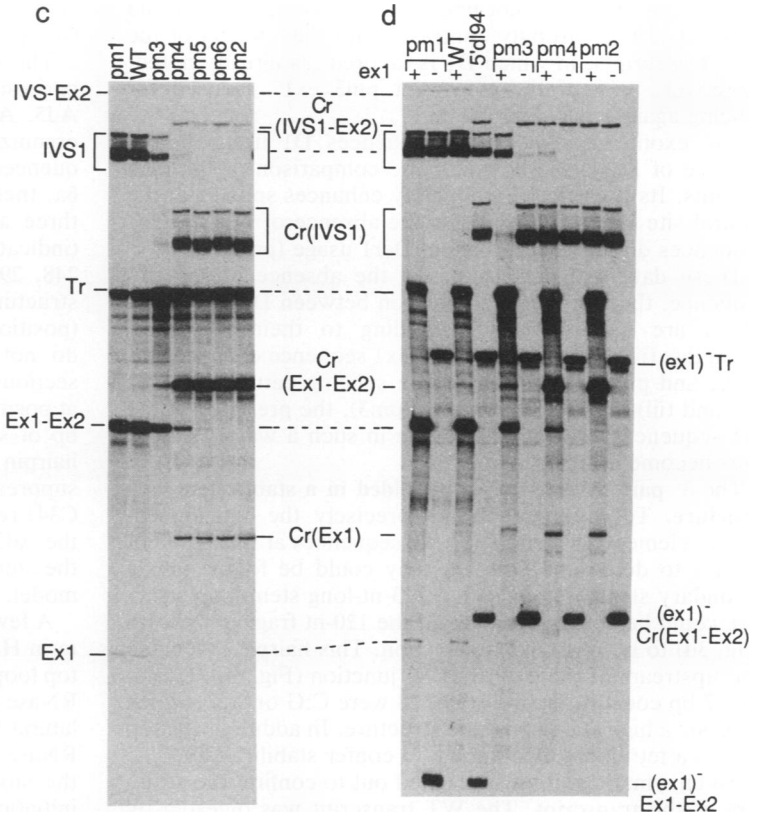

FIG. 4. Effect of the intrinsic strength and of the exonic sequences on the selection of donor sites. (a) Donor site sequence modifications. The sequences of normal (+), suppressed (−), and modified (arrows) D1 and Dcr1 are indicated. Capital letters represent the nucleotides complementary to U1 RNA. The stability of the donor site-U1 RNA duplexes was calculated as described by Freier et al. (16) except that the free energy increments for paired terminal nucleotides were not included. The stretch used for this estimation is overlined (thin lines, A:U; thick lines, G:C; dots, G:U base pairs). (b) List of mutant transcripts and splicing efficiencies. The D1 and Dcr1 composition of the transcripts (symbols from panel a) and the relative strength of the sites (ratio Dcr1/D1 of duplex stabilities) are indicated. Splicing efficiencies were calculated from data for gels similar to those shown in panels c and d. They represent the averages of two independent experiments, each in duplicate. ex1$^+$ and ex1$^-$ indicate the presence or absence of the sequence from nt 117 to 362, deleted in 5′dl 94 (Fig. 3c). (c and d) In vitro splicing. The transcripts listed in panel b were spliced in vitro, and the products were analyzed by gel electrophoresis. The bands representing the intermediates and products of splicing are indicated: on the left, for splicing at D1 of WT transcript or its derivatives (Tr); between the gels, for splicing at Dcr1 of the WT transcript and its derivatives; on the right, for splicing at D1 of 5′dl 94 or its derivatives (ex1$^-$Tr).

suggest that length is not the only exonic parameter involved in the selection.

The relative strength of the two donor sites also affects selection. To determine the influence of the intrinsic strength of the sites in selection, we constructed mutants in which the strength of D1 or Dcr1 or both was modified. The sequence modifications and the estimated stability of the modified donor site-U1RNA duplexes are shown in Fig. 4a, and the list of mutants is presented in Fig. 4b.

The mutants (which have the exact sequence context of the WT transcripts) were assayed for splicing in vitro (Fig. 4c). As expected, the usage of D1 decreases and that of Dcr1 increases upon increase of the Dcr1/D1 strength ratio (Fig. 4b). However, the products of splicing at D1 are still detected when D1 is two to three times weaker than Dcr1 (e.g., in mutants pm4 and pm5). In contrast, the products of splicing at Dcr1 are not detected even when Dcr1 has about the same strength as D1 (WT) and are barely detectable when Dcr1 is 1.4 times stronger than D1 (pm3). This latter transcript is a special case, since only 39% of the precursor is used for splicing (D1 plus Dcr1) against 63 to 81% for all other transcripts, a fact that was observed in three independent experiments.

Thus, the relative strength of the sites is an important parameter of selection of donor sites confirming a *cis* competition between the two sites. However, D1 is clearly preferred to Dcr1, possibly in relation to the presence of the exonic upstream sequence.

The upstream exonic sequence enhances splicing at the natural site. According to the data of Fig. 3, a likely candidate for explaining the preference for D1 in the *cis* competition for U1snRNP is the sequence of 120 nt upstream of the exon 1-IVS1 junction. To test this possibility, we prepared transcripts derived from the deletion mutant 5′dl 94 (Fig. 3) and harboring the same sets of D1 and Dcr1 sequences as pm1, pm2, pm3, and pm4. These transcripts are designated ex1$^-$ to distinguish them from those (ex1$^+$) derived from the WT transcript. They lack most of the exonic 120-nt sequence shown above to be involved in selection.

The splicing patterns of the ex1$^-$ and the ex1$^+$ transcripts were compared. As shown in Fig. 4d, we observe that the presence or absence of the exonic sequence has a considerable influence on the selection of splice sites (compare free IVS1 on the gel and splicing efficiencies in Fig. 4b). D1 and Dcr1 are used roughly according to their relative strengths

(though the relation is not linear) in 5'dl 94, pm3 ex1$^-$, and pm4 ex1$^-$. This strongly contrasts with the splicing of the ex1$^+$ transcripts, in which D1 is favored, as already noted. Moreover, 71% of the transcript pm3 ex1$^-$ is used for splicing against 39% for pm3 ex1$^+$.

The exonic sequence also influences D1 usage in the absence of Dcr1, as shown by the comparison of the pm1 mutants. Its presence reproducibly enhances splicing at the natural site D1. In contrast, in the absence of D1, the ex1 sequences do not seem to affect Dcr1 usage (pm2).

These data indicate that (i) in the absence of the ex1 sequence, there is a *cis* competition between D1 and Dcr1, which are used roughly according to their respective strengths; (ii) the presence of the ex1 sequence enhances D1 usage and provokes a shift of the splicing pattern towards D1; and (iii) in at least one case (pm3), the presence of the ex1 sequence perturbs the system in such a way that both sites become difficult to splice.

The 3' part of exon 1 can be folded in a stable stem-loop structure. To investigate more precisely the splicing enhancer element, we examined the sequences at the 3' end of exon 1 to determine whether they could be folded into a secondary structure. Indeed, a 113-nt-long stem-loop structure could be constructed within the 120-nt fragment shown (Fig. 3d) to be involved in selection. This hairpin is located 6 nt upstream of the exon 1-IVS1 junction (Fig. 5b). Out of the 37 bp constituting the stem, 28 were C:G or G:C, which suggests a high stability of the structure. In addition, the top loop is a tetraloop, also known to confer stability (20).

An enzymatic study was carried out to confirm the structure of the transcript. The WT transcript was digested by two enzymes specific for single-stranded RNA (RNase T1 and micrococcal nuclease) and V1 RNase, specific for double-stranded or stacked regions. The sites of cleavage were mapped by reverse transcriptase extension of an oligonucleotide complementary to a sequence downstream of the region to be studied (Fig. 5a). Sequencing of the cDNA in this region revealed three additional nucleotides relative to the sequence as determined previously (23). They are marked by asterisks in Fig. 5b and designated G339a, C339b, and C339c.

The enzymatic study essentially confirms the existence of the stem-loop structure. A typical analysis is shown in Fig. 5a, and the results of several experiments are reported in Fig. 5b. The single-stranded specific nucleases attack only unpaired nucleotides, whereas V1 nuclease attacks the various segments of the stem. The most frequent and strongest V1 hydrolyses occur in the 3' part of the stem, suggesting that these regions are the most accessible to the enzyme in the tertiary RNA configuration.

A few discrepancies were noted. First, V1 RNase hydrolyzes the top part of the internal loops between stems III and IV and between stems VI and VII as well as nucleotides at the foot of the hairpin. This may be attributed to either the existence of noncanonical hydrogen bonds (A:G in loop III-IV, for instance) or stacking in these regions. Second, the nucleotides around the exon-intron junction do not seem to be attacked by either single-stranded or double-stranded nucleases. This is the consequence of a pause of the reverse transcriptase at the foot of the hairpin which precludes the detection of cleavage at these sites (3, 15). It may also be remarked that the stability of the stem of the lateral hairpin is rather low (one C:G and one G:U base pair). It is nevertheless shown as a hairpin rather than a loop for two reasons. First, there is no single-stranded nuclease attack at the level of the 2 bp; second, the stem is reinforced by an additional base pair in the Ad5 homologous structure (Fig. 6a; see below).

The stem-loop structure was further confirmed by comparison with the homologous sequence of region E3 from Ad5. Ad5 is the only adenovirus with an E3 exon-intron organization similar to that of Ad2, which has been sequenced (9). In the stem-loop structure region shown in Fig. 6a, there are 13 base substitutions and the deletion of the three additional nucleotides GCC that we found in Ad2 (indicated by dashes in the figure). Substitutions at positions 248, 295, 325, 367, and 383 do not modify the stem-loop structure. Two pairs of substitutions are compensatory (positions 263 and 266; positions 332 and 334), that is, they do not change the number of base pairs (the modified sections are shown in the insets of Fig. 6a). The substitutions at positions 294 and 296 and the deletion of GCC suppress 1 bp of stem section IV and add an A:U pair in the lateral hairpin (inset). Finally, the substitution at position 313 suppresses 1 bp of stem section I and the substitution of C341 replaces a G:C by a G:U base pair. The comparison of the Ad2 and Ad5 structures shows that the main features of the stem-loop structure are conserved, which supports our model.

A few characteristics of the stem-loop structure which we term HP1 should be outlined (Fig. 6b). The sequence of the top loop is ACUG, the last G of which is very sensitive to T1 RNase attack. The same sequence is found in the loop of the lateral hairpin, where its last G is also very sensitive to T1 RNase. In addition, the loop of the lateral hairpin harbors the stop codon UAG of the late L4 protein whereas the initiation codon AUG of a putative early protein is partially included in the stem (Fig. 1a). Furthermore, there are two lateral loops on the 3' side made only of A and C (AC loops; dotted lines in Fig. 6b). All these features are conserved in Ad5. Their exact meaning is still obscure but will have to be taken into consideration in further functional investigations.

Thus, at least in the WT transcript in solution, the 3' end of exon 1 is folded in a stem-loop structure. The data suggest that HP1 is the D1 enhancer element: (i) its 113 nt are entirely contained in the 120-nt stretch shown to be required for the selection of D1 and silencing of Dcr1 (Fig. 3d); (ii) HP1 is present in both the WT and 5'dl 41 transcripts, which select D1 and exclude Dcr1; and (iii) HP1 is disrupted in transcripts 5'dl 59, 5'dl 62, and 5'dl 94 (Fig. 6b), which are permissive for both D1 and Dcr1.

DISCUSSION

As a consequence of the variability of the donor site sequence (26), the existence of numerous potential donor sites in primary transcripts can be predicted. Some of them are used, constitutively or alternatively, for the formation of mRNAs. Other sites, normally silent, are revealed upon mutation of the natural sites and constitute the cryptic sites. The cryptic donor site sequences are mostly identical to those of natural sites and are intrinsically functional. Therefore, there must be some mechanism to prevent their use. Such a mechanism is expected to be a general one and to require general *trans*-acting factors since it ought to occur for many gene transcripts, independently of tissue or the stage of life cycle. To study this mechanism, we searched for *cis*-acting sequences involved in the prevention of a cryptic site, Dcr1, located 74 nt downstream of the first natural donor site D1 of the E3 transcription unit of Ad2.

Three *cis*-acting elements are involved in splice donor site selection. Our data show that there are at least three *cis*-

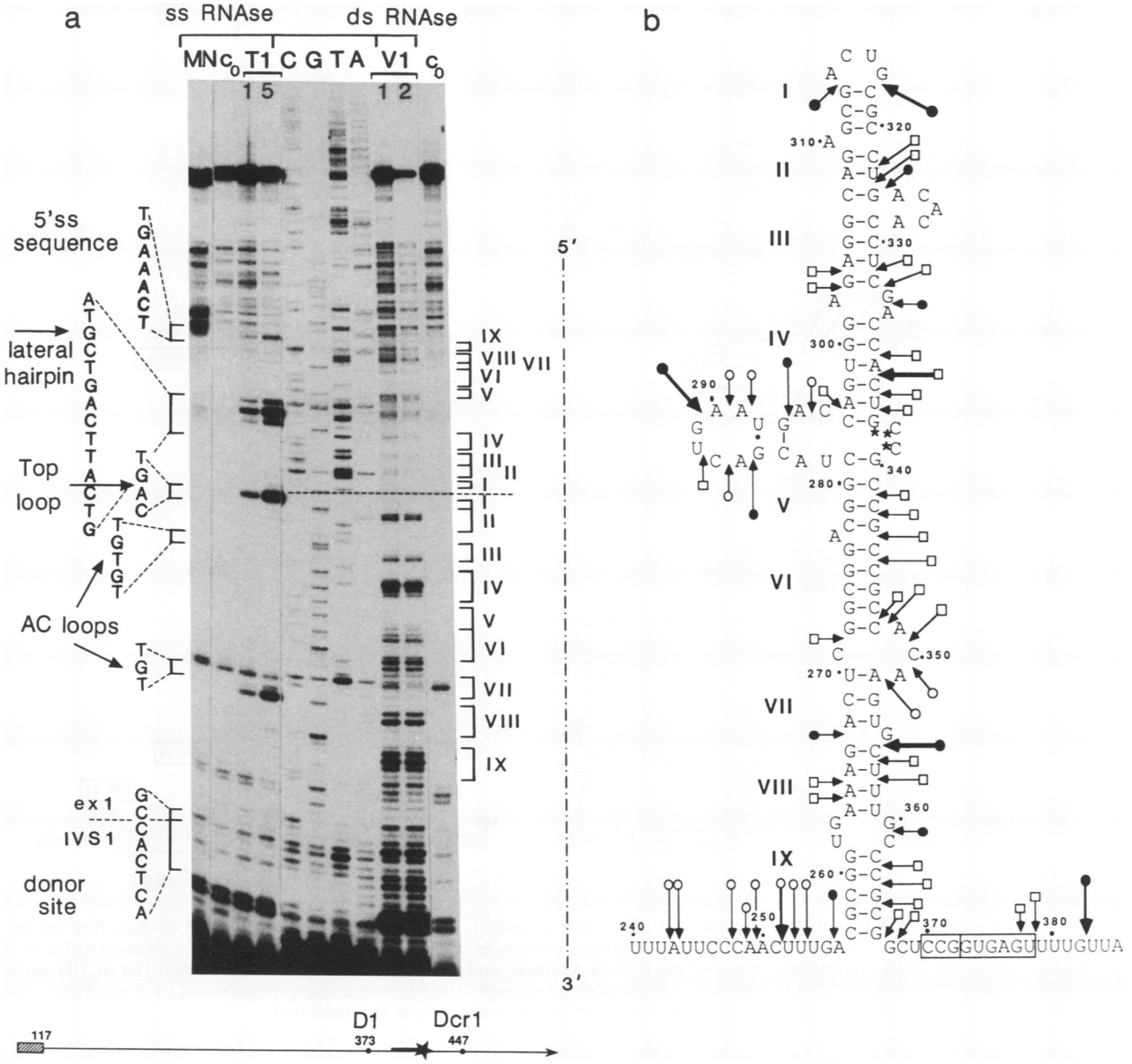

FIG. 5. Stem-loop structure at the 3′ end of E3 exon 1. The WT transcript was treated by RNases specific for single-stranded (ss) or double-stranded (ds) RNA. The sites of hydrolysis were determined by extension of a 5′ end-labeled primer 5′-GGCCCTCGATATGCTCT TCGGGCAATTC-3′ complementary to the sequence between positions 417 and 390 (bottom) and comparison with a sequencing ladder (lanes C, G, T, A in panel a) made by reverse transcription of the corresponding DNA fragment. A representative gel is shown in panel a. Sequences are those of the coding strand. Other enzyme concentrations and running times were used to extend the various regions. The results are summarized in panel b. (a) MN, T1, and V1 refer to micrococcal, T1, and V1 nucleases respectively and c_0 to a control treated under the same conditions but in the absence of enzymes. 1, standard enzyme concentration (Materials and Methods); 2 and 5, 2× and 5× concentrations, respectively. (b) The arrows indicate the sites of nuclease attack. Their importance varies according to attack intensity. Nucleases: ○→, micrococcal; ●→, T1; □→, V1. The different stem fragments are numbered I to IX (see panel a). The donor site is boxed. The asterisks point to the three additional nucleotides as discussed in the text.

acting elements which participate in the selection of the natural and cryptic donor sites in this system.

The first two are the two donor site sequences themselves. When the first intronic G of the natural site D1 is mutated (pm2), when the D1 sequence is deleted (5′dl 37, 5′dl 36, 5′dl 67), or when other mutations weaken D1 relative to Dcr1 (pm4), then the cryptic site is activated. This is in general agreement with the data obtained in studies of other genes after spontaneous mutations at G1, G(−1), or G5 (24, 29, 37, 44, 56, 58) or after site-directed mutagenesis (2, 61). Changes in the Dcr1 sequences which increase its strength (pm3) also allow Dcr1 usage. In fact, it is the relative strength of the two sites which should be considered. It has as a consequence a *cis* competition between the sites which is primarily if not exclusively dependent on their relative strength, when the third element is deleted (transcripts derived from 5′dl 94).

The third element is located in the 120-nt stretch at the 3′ extremity of exon 1 and was revealed by mutants 5′dl 59,

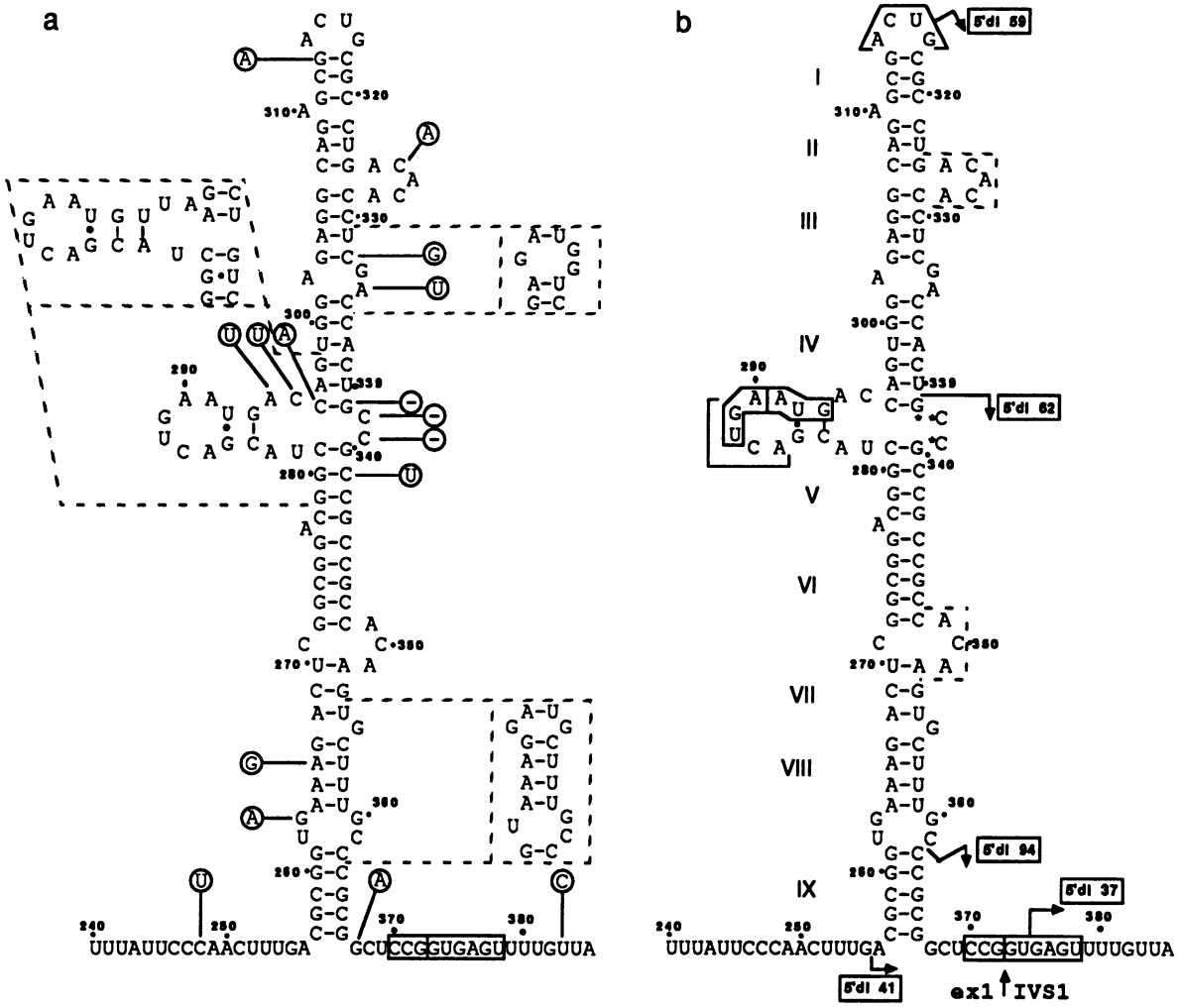

FIG. 6. (a) Comparison of Ad2 and Ad5 structures at the 3′ end of exon 1. The stem-loop structure is that of Ad2. The nucleotide substitutions or deletions (−) in Ad5 are circled. The local structural changes in Ad5 are shown in insets. (b) Some characteristics of HP1. The stop codon of protein pVIII, the initiation codon of the E3A-12.5 protein, and the donor site D1 (Fig. 1) are boxed. The ACUG sequence found in the top loop of HP1 and repeated in the lateral hairpin loop is overlined. The AC lateral loops are indicated by dashed lines. Notice that these sequences are all conserved in Ad5. The start of deletion mutant transcripts is also indicated.

5′dl 62, and 5′dl 94. Its presence modifies the results of the *cis* competition between the two sites (Fig. 4d). This exonic upstream element enhances splicing at D1 (pm1) in such a way that Dcr1 has to become much stronger than D1 in order to be used (pm3, pm4). Thus, the actual strength of a site may be higher than its intrinsic strength (i.e., its capacity to bind U1RNA), owing to the presence of an exonic enhancer element, and this may finally lead to the silencing of the competing site.

Thus, for the silencing of a cryptic site, the sequence context is as important as it is for alternative splicing (10, 11, 17, 21, 27, 31, 33, 38, 43, 45, 47, 52, 54, 57). In our system, the balance due to the simple *cis*-competition between the sites is modified by the presence of an exonic fragment which has an enhancing activity on the proximal donor site. This mechanism might be a general one, though the nature of the sequence context could vary according to the system.

Characteristics of the upstream exonic element. The 120-nt segment which has an enhancing effect on splicing at D1 can be folded in a 113-nt-long hairpin (HP1) punctuated by bulges and internal loops and by a lateral weak hairpin. The hairpin structure was confirmed by a phylogenetic comparison and by an experimental enzymatic study of the transcript in solution. Its stability is primarily due to a large number of G:C or C:G base pairs and to a top tetraloop. Although this has not yet been proven experimentally, the stability of HP1 strongly suggests that it may be conserved upon addition of the nuclear factors which will ultimately give rise to active spliceosomes, as is the case for other RNAs (small nuclear RNA, rRNA) upon formation of their respective ribonucleoproteins.

The large size of HP1, intermediate between that of U5 and U6 small nuclear RNAs, is compatible with the binding of several *trans*-acting factors and with multiple functions. In addition to the enhancing effect of its sequence on splicing at the proximal donor site, HP1 may also be involved in the regulation of protein synthesis, as suggested by the presence, in its lateral hairpin, of the stop codon for the late protein pVIII and of the initiation codon for the putative early 12.5K protein (Fig. 6b).

Secondary structure or conformation of the pre-mRNA has often been invoked as a possible element for splicing regulation, and data in favor of this proposition are starting to accumulate. Artificial hairpins inserted in a premRNA may induce alternative splicing under determined conditions (15, 51). The formation of a small natural hairpin in the E1A transcript of Ad2 is able to bring otherwise distant branch sites to an operational distance of the acceptor site and thus allow the excision of an optional intron (7). The sequestration of the 3′ splice site of the C4-M1 intron from the immunoglobin μ heavy-chain pre-mRNA in a stem-loop structure inhibits the splicing reaction and thus participates in the selection between the membrane-bound and secreted forms of the immunoglobin (59). A complex stem-loop structure was built in the region of the optional mutually exclusive exons 6A and 6B of the β tropomyosin gene transcript; mutational studies showed that this structure is involved in the particular characteristics of splicing of the two exons (31). Though the function of the *rev* protein from immunodeficiency virus in splicing regulation is still controversial (6, 32, 42), it may also be stressed that its *cis*-acting response element resides in a stem-loop structure (22, 32, 40). In the case of the Ad2 E3 transcript, it is again a region with a potential stem-loop structure (HP1) which is involved in the selection of splice sites.

ACKNOWLEDGMENTS

We thank C. Ehresmann and her group for competent advice in secondary structure determination and fruitful discussion and P. Simpson for critically reading the manuscript. We are grateful to the cell culture staff for providing HeLa cells, to A. Staub for oligodeoxynucleotide synthesis, to C. Werlé and B. Boulay for the preparation of illustrations, and to the secretarial staff for typing the manuscript.

This work was supported by grants from the Centre National de la Recherche Scientifique, from the Institut National de la Santé et de la Recherche Médicale, from the Association de la Recherche sur le Cancer, and from the Fondation pour la Recherche Médicale.

REFERENCES

1. **Adami, G. R., C. W. Marlor, N. L. Barrett, and G. G. Carmichael.** 1989. Leader-to-leader splicing is required for efficient production and accumulation of polyomavirus late mRNAs. J. Virol. **63**:85–93.
2. **Aebi, M., H. Hornig, R. A. Padgett, J. Reiser, and C. Weissmann.** 1986. Sequence requirements for splicing of higher eukaryotic nuclear pre-mRNA. Cell **47**:555–565.
3. **Baudin, F., C. Ehresmann, P. Romby, M. Mougel, J. Colin, L. Lempereur, J.-P. Bachellerie, J.-P. Ebel, and B. Ehresmann.** 1987. Higher-order structure of domain III in Escherichia coli 16S ribosomal RNA, 30S subunit and 70S ribosome. Biochimie **69**:1081–1096.
4. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
5. **Breitbart, R. E., A. Andreadis, and B. Nadal-Ginard.** 1987. Alternative splicing: a ubiquitous mechanism for the generation of multiple protein isoforms from single genes. Annu. Rev. Biochem. **56**:467–495.
6. **Chang, D. D., and P. A. Sharp.** 1989. Regulation by HIV rev depends upon recognition of splice sites. Cell **59**:789–795.
7. **Chebli, K., R. Gattoni, P. Schmitt, G. Hildwein, and J. Stevenin.** 1989. The 216-nucleotide intron of the EIA pre-mRNA contains a hairpin structure that permits utilization of unusually distant branch acceptors. Mol. Cell. Biol. **9**:4852–4861.
8. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**:1023–1030.
9. **Cladaras, C., and W. S. M. Wold.** 1985. DNA sequence of the early E3 transcription unit of adenovirus 5. Virology **140**:28–43.
10. **Cooper, T. A., M. H. Cardone, and C. P. Ordahl.** 1988. Cis requirements for alternative splicing of the cardiac troponin T pre-mRNA. Nucleic Acids Res. **16**:8443–8465.
11. **Cooper, T. A., and C. P. Ordahl.** 1989. Nucleotide substitutions within the cardiac troponin T alternative exon disrupt pre-mRNA alternative splicing. Nucleic Acids Res. **17**:7905–7921.
12. **Dignam, J. D., R. M. Lebovitz, and R. G. Roeder.** 1983. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res. **11**:1475–1489.
13. **Ehresmann, C., F. Baudin, M. Mougel, P. Romby, J.-P. Ebel, and B. Ehresmann.** 1987. Probing the structure of RNAs in solution. Nucleic Acids Res. **15**:9109–9128.
14. **Eperon, L. P., J. P. Estibeiro, and I. C. Eperon.** 1986. The role of nucleotide sequences in splice site selection in eukaryotic pre-messenger RNA. Nature (London) **324**:280–282.
15. **Eperon, L. P., I. R. Graham, A. D. Griffiths, and I. C. Eperon.** 1988. Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? Cell **54**:393–401.
16. **Freier, S. M., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. H. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Biochemistry **83**:9373–9377.
17. **Gallego, M. E., and B. Nadal-Ginard.** 1990. Myosin light-chain 1/3 gene alternative splicing: *cis* regulation is based upon a hierarchical compatibility between splice sites. Mol. Cell. Biol. **10**:2133–2144.
18. **Gallinaro, H., P. Vincendon, A. Sittler, and M. Jacob.** 1988. Alternative use of a polyadenylation signal and of a downstream 3′ splice site. Effect of 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole. J. Mol. Biol. **204**:1031–1040.
19. **Green, M. R., T. Maniatis, and D. A. Melton.** 1983. Human β-globin pre-mRNA synthesized in vitro is accurately spliced in Xenopus oocyte nuclei. Cell **32**:681–694.
20. **Groebe, D. R., and O. C. Uhlenbeck.** 1988. Characterization of RNA hairpin loop stability. Nucleic Acids Res. **16**:11725–11735.
21. **Hampson, R. K., L. La Follette, and F. M. Rottman.** 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. Mol. Cell. Biol. **9**:1604–1610.
22. **Heaphy, S., C. Dingwall, I. Ernberg, M. J. Gait, S. M. Green, J. Karn, A. D. Lowe, M. Singh, and M. A. Skinner.** 1990. HIV-1 regulator of virion expression (Rev) protein binds to an RNA stem-loop structure located within the rev response element region. Cell **60**:685–693.
23. **Hérissé, J., G. Courtois, and F. Galibert.** 1980. Nucleotide sequence of the EcoRI D fragment of adenovirus 2 genome. Nucleic Acids Res. **8**:2173–2192.
24. **Hodges, P. E., and L. E. Rosenberg.** 1989. The spf^{ash} mouse: a missense mutation in the ornithine transcarbamylase gene also causes aberrant mRNA splicing. Proc. Natl. Acad. Sci. USA **86**:4142–4146.
25. **Inouye, S., and M. Inouye.** 1987. Oligonucleotide-directed site-specific mutagenesis using double-stranded plasmid DNA, p. 181–206. *In* S. A. Narang (ed.), Synthesis and applications of DNA and RNA. Academic Press, Inc., Orlando, Fla.
26. **Jacob, M., and H. Gallinaro.** 1989. The 5′ splice site: phylogenetic evolution and variable geometry of association with U1RNA. Nucleic Acids Res. **17**:2159–2180.
27. **Kakizuka, A., T. Ingi, T. Murai, and S. Nakanishi.** 1990. A set of U1 snRNA-complementary sequences involved in governing alternative RNA splicing of the kininogen genes. J. Biol. Chem. **265**:10102–10108.
28. **Krainer, A. R., T. Maniatis, B. Ruskin, and M. R. Green.** 1984. Normal and mutant human β-globin pre-mRNAs are faithfully and efficiently spliced in vitro. Cell **36**:993–1005.
29. **Lapoumeroulie, C., S. Acuto, F. Rouabhi, D. Labie, R. Krishnamoorthy, and A. Bank.** 1987. Expression of a β thalassemia gene with abnormal splicing. Nucleic Acids Res. **15**:8195–8204.
30. **Latchman, D. S.** 1990. Cell-type-specific splicing factors and the

regulation of alternative RNA splicing. New Biologist **2**:297–303.
31. **Libri, D., M. Goux-Pelletan, E. Brody, and M. Y. Fiszman.** 1990. Exon as well as intron sequences are *cis*-regulating elements for the mutually exclusive alternative splicing of the β tropomyosin gene. Mol. Cell. Biol. **10**:5036–5046.
32. **Malim, M. H., J. Hauber, S.-Y. Le, J. V. Maizel, and B. R. Cullen.** 1989. The HIV-1 rev trans-activator acts through a structured target sequence to activate nuclear export of unspliced viral mRNA. Nature (London) **338**:254–257.
33. **Mardon, H. J., G. Sebastio, and F. E. Baralle.** 1987. A role for exon sequences in alternative splicing of the human fibronectin gene. Nucleic Acids Res. **15**:7725–7733.
34. **Mayeda, A., and Y. Ohshima.** 1990. β-Globin transcripts carrying a single intron with three adjacent nucleotides of 5′ exon are efficiently spliced in vitro irrespective of intron position or surrounding exon sequences. Nucleic Acids Res. **18**:4671–4676.
35. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
36. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
37. **Nakajima, H., N. Kono, T. Yamasaki, K. Hotta, M. Kawachi, M. Kuwajima, T. Noguchi, T. Tanaka, and S. Tarui.** 1990. Genetic defect in muscle phosphofructokinase deficiency. Abnormal splicing of the muscle phosphofructokinase gene due to a point mutation at the 5′-splice site. J. Biol. Chem. **265**:9392–9395.
38. **Nasim, F. H., P. A. Spears, H. M. Hoffmann, H.-C. Kuo, and P. J. Grabowski.** 1990. A sequential splicing mechanism promotes selection of an optimal exon by repositioning a downstream 5′ splice site in preprotachykinin pre-mRNA. Genes Dev. **4**:1172–1184.
39. **Nelson, K. K., and M. R. Green.** 1990. Mechanism for cryptic splice site activation during pre-mRNA splicing. Proc. Natl. Acad. Sci. USA **87**:6253–6257.
40. **Olsen, H. S., P. Nelbock, A. W. Cochrane, and C. A. Rosen.** 1990. Secondary structure is the major determinant for interaction of HIV rev protein with RNA. Science **247**:845–848.
41. **Parent, A., S. Zeitlin, and A. Efstratiadis.** 1987. Minimal exon sequence requirements for efficient in vitro splicing of monointronic nuclear pre-mRNA. J. Biol. Chem. **262**:11284–11291.
42. **Pavlakis, G. N., and B. K. Felber.** 1990. Regulation of expression of human immunodeficiency virus. New Biologist **2**:20–31.
43. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice-site selection. Cell **46**:681–690.
44. **Ris-Stalpers, C., G. G. J. M. Kuiper, P. W. Faber, H. U. Schweikert, H. C. J. Van Rooij, N. D. Zegers, M. B. Hodgins, H. J. Degenhart, J. Trapman, and A. O. Brinkmann.** 1990. Aberrant splicing of androgen receptor mRNA results in synthesis of a nonfunctional receptor protein in a patient with androgen insensitivity. Proc. Natl. Acad. Sci. USA **87**:7866–7870.
45. **Saga, Y., J. S. Lee, C. Saraiya, and E. A. Boyse.** 1990. Regulation of alternative splicing in the generation of isoforms of the mouse Ly-5 (CD45) glycoprotein. Proc. Natl. Acad. Sci. USA **87**:3728–3732.
46. **Schmitt, P., R. Gattoni, P. Keohavong, and J. Stevenin.** 1987. Alternative splicing of E1A transcripts of adenovirus requires appropriate ionic conditions in vitro. Cell **50**:31–39.
47. **Siebel, C. W., and D. C. Rio.** 1990. Regulated splicing of the Drosophila P transposable element third intron in vitro: somatic repression. Science **248**:1200–1208.
48. **Sittler, A., H. Gallinaro, and M. Jacob.** 1986. In vivo splicing of the premRNAs from early region 3 of adenovirus-2: the products of cleavage at the 5′ splice site of the common intron. Nucleic Acids Res. **14**:1187–1207.
49. **Sittler, A., H. Gallinaro, L. Kister, and M. Jacob.** 1987. In vivo degradation pathway of an excised intervening sequence. J. Mol. Biol. **197**:737–742.
50. **Smith, C. W. J., J. G. Patton, and B. Nadal-Ginard.** 1989. Alternative splicing in the control of gene expression. Annu. Rev. Genet. **23**:527–577.
51. **Solnick, D., and S. I. Lee.** 1987. Amount of RNA secondary structure required to induce an alternative splice. Mol. Cell. Biol. **7**:3194–3198.
52. **Somasekhar, M. B., and J. E. Mertz.** 1985. Exon mutations that affect the choice of splice sites used in processing the SV40 late transcripts. Nucleic Acids Res. **13**:5591–5609.
53. **Stalhandske, P., H. Persson, M. Perricaudet, L. Philipson, and U. Petersson.** 1983. Structure of three spliced mRNAs from region E3 of adenovirus type 2. Gene **22**:157–165.
54. **Streuli, M., and H. Saito.** 1989. Regulation of tissue-specific alternative splicing exon-specific cis-elements govern the splicing of leukocyte common antigen pre-mRNA. EMBO J. **8**:787–796.
55. **Tollefson, A. E., P. Krajcsi, S. Yei, C. R. Carlin, and W. S. M. Wold.** 1990. A 10,400-molecular-weight membrane protein is coded by region E3 of adenovirus. J. Virol. **64**:794–801.
56. **Treisman, R., S. H. Orkin, and T. Maniatis.** 1983. Specific transcription and RNA splicing defects in five cloned β-thalassaemia genes. Nature (London) **302**:591–596.
57. **Tsai, A. Y. M., M. Streuli, and H. Saito.** 1989. Integrity of the exon 6 sequence is essential for tissue-specific alternative splicing of human leukocyte common antigen pre-mRNA. Mol. Cell. Biol. **9**:4550–4555.
58. **Vidaud, M., R. Gattoni, J. Stevenin, D. Vidaud, S. Amselem, J. Chibani, J. Rosa, and M. Goossens.** 1989. A 5′ splice-region G→C mutation in exon 1 of the human β-globin gene inhibits pre-mRNA splicing: a mechanism for β^+-thalassemia. Proc. Natl. Acad. Sci. USA **86**:1041–1045.
59. **Watakabe, A., K. Inoue, H. Sakamoto, and Y. Shimura.** 1989. A secondary structure at the 3′ splice site affects the in vitro splicing reaction of mouse immunoglobulin μ chain pre-mRNAs. Nucleic Acids Res. **17**:8159–8169.
60. **Weber, S., and M. Aebi.** 1988. In vitro splicing of mRNA precursors: 5′ cleavage site can be predicted from the interaction between the 5′ splice region and the 5′ terminus of U1 snRNA. Nucleic Acids Res. **16**:471–486.
61. **Wieringa, B., F. Meyer, J. Reiser, and C. Weissmann.** 1983. Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β-globin gene. Nature (London) **301**:38–43.
62. **Winship, P. R.** 1989. An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. Nucleic Acids Res. **17**:1266.