A cis-Acting Selector of a 5' Splice Site

COOPERATION BETWEEN THE SEQUENCE OF THE SITE AND AN UPSTREAM EXONIC ELEMENT*

(Received for publication, March 21, 1993, and in revised form, June 21, 1993)

Liliane Kister, Lionel Domenjoud, Hélène Gallinaro, and Jacob Monique

From the 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 Cedex, France

To investigate the mechanism by which a 5' splice site (D1) is selected, while a nearby potentially functional site (Dcr1) is silenced, we have studied the importance of the 9-nucleotide sequence of these 5' splice sites for their respective usage. Our model system uses a transcript derived from the early transcription unit 3 of adenovirus-2. Transcripts, harboring an exonic element previously shown to be required for the selection of D1 in the presence of Dcr1, were mutated in the D1 and Dcr1 sequences and assayed for splicing in vitro. We first show that an increased ability of D1 to pair with U1 small nuclear (sn) RNA correlates with an increased accumulation of splicing intermediates, independently of the presence of Dcr1. This variation of efficiency of the first splicing reaction does not significantly affect the overall splicing efficiency except when the potential D1-U1 snRNA hybrid is less than 6 base pairs. Equally, the selector activity of the upstream exon element requires a D1 sequence that is able to form hybrids of 6 base pairs or more with U1 snRNA. This indicates that the cis-acting selector of D1 includes the exonic element (a potential stem-loop structure) and a D1 sequence of sufficient strength.

The presence of multiple introns in many mammalian genes allows the formation of several mRNAs and proteins from a single gene. The introns may be excised, or not, as a consequence of alternative splicing pathways, which may be stage or tissue specific (for review see Breithart et al., 1987; Smith et al., 1989; Latchman, 1990; McKeown, 1990; Rio, 1992a, 1992b; Mattox et al., 1992). The choice of a given splicing pathway implies a selection of splice sites. The search for the mechanisms underlying this selection has become one of the key problems in understanding gene expression. Consensus sequences have been found at the two extremities of the introns, thus defining a 5' splice site (or donor site) and a 3' splice site (or acceptor site). Depending on the gene and the pathway, selection may concern both sites, or only one of them, and complex combinations are often observed in genes harboring several introns. Various cis-acting elements and trans-acting factors have been shown to participate in the selection.

For the elucidation of one of these mechanisms, we have chosen a simple case: the exclusive selection of a 5' splice site in the presence of a nearby, potentially functional, but normally silent cryptic site. The cryptic site can be induced when the natural site is suppressed and the cryptic intron can be excised using the same 5' splice site as the natural intron (Domenjoud et al., 1991). Such a situation is relatively common as shown by the discovery of cryptic sites in various genes after sequence modifications, including mutations leading to genetic diseases. The frequent occurrence of cryptic sites is primarily the consequence of the degeneracy of the 9-nucleotide-long 5' splice site sequence (Jacob and Gallinaro, 1989), as random sequences often resemble a 5' splice site.

Our model system derives from an early transcription unit (E3) from adenovirus-2. In the natural context, the 5' splice site of the first intron, designated D1, is used for the formation of all E3 mRNAs, whereas the cryptic site, Dcr1, located 74 nucleotides downstream is silent (Domenjoud et al., 1991). Some of the parameters involved in the exclusive selection of D1 and the silencing of Dcr1 have been defined (Domenjoud et al., 1991, 1993). First, the kinetics of splicing of the two sites are quite different; in particular early splicesome assembly at Dcr1 is slower relative to assembly at D1. Second, a potential stem-loop structure located close to D1 in the upstream exon protects D1 again cis-competition by Dcr1 and therefore acts as a selector of D1. Third, the relative intrinsic strengths of D1 and Dcr1 also play a role in the selection. Intrinsic strength is defined as the capacity of a 5' splice site to pair to the 5' invariant extremity of U1 snRNA. We demonstrated that some changes of the intrinsic strengths of the sites allowed Dcr1 usage in the presence of D1. However, all the data could not be explained on this basis (Domenjoud et al., 1991), and we have undertaken a more detailed analysis of the effect of the intrinsic strength of the sites on selection.

Here, we show that the intrinsic strength of the D1 site determines the amount of intermediates formed after the first splicing reaction, without necessarily affecting the overall splicing efficiency. We also report that a D1 sequence with the capacity to pair stably with U1 snRNA is absolutely required for the selector activity of the upstream exon element. This indicates that the cis-acting selector of D1 includes the sequence of D1 in addition to the previously described exonic element.

* This work was supported by grants from the Centre National de la Recherche Scientifique, the Institut National de la Santé et de la Recherche Médicale, the Centre Hospitalier Universitaire Régional, the Fondation pour la Recherche Médicale, and the Association pour la Recherche contre le Cancer. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Present address: Laboratoire d'Enzymologie et de Génie Génétique, Boulevard des Aiguillettes, 54000 Vandoeuvre les Nancy, France.

§ To whom correspondence should be addressed. Tel: 33-88-37-12-56; Fax: 33-88-37-01-48.

1 The abbreviations used are: snRNA, small nuclear RNA; bp, base pair(s); IVS, intervening sequence.
FIG. 1. The effect of intrinsic strength of D1 on splicing efficiency and induction of the cryptic site. A, schematic representation of the wild type transcript. The position of the sites of interest is indicated. PL indicates polylinker sequences. Numbering is according to (Cladaras and Wold, 1985). The intron-exon organization is shown below. The structure of the HP1-D1 region is shown in Fig. 3. B and C, Dcr1- and Dcr1+ transcripts were studied in parallel. The various D1 sequences are given in Table I. The transcripts were spliced in vitro under standard conditions. Two typical gels are shown. The intermediates and products of splicing at

MATERIALS AND METHODS

Construction of Mutants—All DNA fragments to be transcribed were inserted in the polylinker of plasmid pSP64 (Melton et al., 1984). The wild type fragment was a PstI-SacI fragment of 846 bp from the proximal part of the adenovirus-2 (Ad2) early region 3 transcription unit (E3) harboring a natural and a cryptic 5' splice site (D1 and Dcr1, respectively, Fig. 1A), as described previously (Domenjoud et al., 1991).

Only mutants where the sequence of the 5' splice sites D1 and/or Dcr1 were modified were used for the present work. They were prepared from the wild type transcript by site-directed mutagenesis (Inouye and Inouye, 1987). The preparation of mutants pm0 to pm6 has been described (Domenjoud et al., 1991). The sequence of D1 and Dcr1 in all mutants is listed in Table I. A pm1 construction (natural D1 site, Dcr1 suppressed by substitution of the first G of the cryptic intron by an A) was used for the preparation of the Dcr1− transcripts (mutants pm12, pm18, pm8, pm16, pm14, pm7, and pm10) by site-directed mutagenesis. Mutants pm13, pm19, pm9, pm17, pm15, and pm11 (Dcr1+ transcripts) were similarly prepared from the wild type transcript (pm0). This series of mutants, designated as pm differs only in the sequence of 5' splice sites, which implies, in particular, that exon 1 is intact. Their sequence was determined using a modification of the dideoxynucleotide sequencing procedure (Winship, 1989).

RNA Synthesis and in Vitro Splicing—These were as described previously (Domenjoud et al., 1991, 1993). Briefly, the pSP64 plasmids containing the inserted E3 fragments were linearized by EcoRI and transcribed in the presence of [α-32P]CTP. The specific activity of the transcripts was sufficient to detect splicing intermediates using 100,000 cpm Cerenkov of RNA precursor/assay. The standard conditions for in vitro splicing were as described (Domenjoud et al., 1991).

For determination of splicing efficiency, incubation was for 3 h; for kinetic studies, splicing efficiency was determined after 0, 10, 15, 20, 25, 30, 40, 50, 60, 90, 120, 150, 180, and 240 min of incubation. A single preparation of HeLa cell nuclear extract allowing a splicing efficiency of about 80% for the wild type transcript was used in these experiments. Splicing products were analyzed in 5% polyacrylamide gels (acylamide/bisacrylamide, 20/1, w/w) containing 8 M urea. The bands corresponding to the different molecules as well as appropriate blank gel fragments were cut out and their radioactivity counted. The amount of intermediates (i), products (p), and remaining precursor (pr) at each time was determined. Their proportion relative to the initial precursor, as indicated in the figures is (i + p)/(i + p + pr), or (p × 100)/(i + p + pr), respectively. Splicing efficiency is (i + p) × 100/(i + p + pr). All experimental values are the average of duplicates. To optimize comparisons, transcripts to be compared were labeled using the same CTP batch; they were used fresh (within a week) and spliced in parallel.

RESULTS

Characteristics of the System and Rationale of the Experiments—Our basic transcript for in vitro studies (WT-pm0) is derived from the proximal part of the E3 region of adenovirus-2 (Domenjoud et al., 1991) and includes only natural contiguous E3 sequences (Fig. 1A). A cryptic 5' splice site, Dcr1, is located 74 nucleotides downstream of the natural site D1. All the cis-acting elements required for an efficient in vitro splicing at D1 and for the silencing of Dcr1 are included in WT-pm0.

The cryptic site Dcr1 can be induced by the suppression of the natural site D1, and the same acceptor site is used for the excision of the natural and cryptic IVS (Domenjoud et al., 1991). We have demonstrated that very different kinetics of splicing at D1 and Dcr1 can partially explain the predominance of D1 over Dcr1 (Domenjoud et al., 1993). In addition, the exclusive selection of D1 requires a cis-acting element located at the 3' extremity of exon 1 (Fig. 1A); this 120-D1 and Dcr1 are displayed as indicated. The free lariat IVS's are found as wide bands; this is due to the trimming of the lariat tail in our in vitro system (Domenjoud et al., 1991). D, splicing efficiency was determined and the means of several experiments (column n) are shown.
nucleotide region can be folded into a potential stem-loop structure designated as HP1 (whose structure is shown in Fig. 3). The deletion of HP1 does not change the kinetics or efficiency of splicing at D1 or Dcr1, when the two sites are considered independently, but the presence of HP1 is required for efficient splicing at D1 in the case of competition by Dcr1; thus, it has a selector activity.

Another element of importance for the selection of a donor site is its capacity to hybridize with the 5' extremity of U1 snRNA, which will be designated as "intrinsic strength" to distinguish it from its global strength, which may involve other characteristics of the site. The number of contiguous base pairs included in the potential hybrid and the free energy increments (–kcal/mol) are indicated for each sequence in Table I. For the sake of simplicity, only the number of base pairs will be indicated in the text and figures, except when otherwise required. D1 and Dcr1 have close intrinsic strengths (pm0, Table I), so that a large difference of intrinsic strengths is not a major requirement for the selection of D1 in the context of our wild type transcript. Nevertheless, we have shown that changes of the relative intrinsic strengths of the sites may allow Dcr1 usage and almost silence D1 (Domenjoud et al., 1991). Some aspects of this effect were intriguing, because the apparent protection of D1 by exon 1 seemed to be lost in certain cases. To understand the underlying mechanisms, we performed a systematic study of the effect of intrinsic strength on the selection of the natural and cryptic splice sites. To this end, plasmids harboring various mutations at the D1 and Dcr1 sites were constructed and used for transcription and in vitro splicing, as will be described below. In all transcripts, the exon 1 sequence was wild type (as well as the other parts of the transcripts), so that the exonic element with the selector activity was present. For clarity, the transcripts will be designated solely by their D1-Dcr1 combination, as listed in Table I.

We have shown previously (Domenjoud et al., 1993) that a kinetic study of splicing could shed some light on the mechanisms underlying changes of splicing efficiency. In particular, the separate analysis of the time courses of accumulation of splicing intermediates and products may provide information on the steps of the reaction at which a given change of the sequence of the transcripts influences splicing. Therefore, when required, kinetic studies were performed to complement the data obtained by the standard determination of splicing efficiency (3 h of reaction).

**The Repercussions of Changes of Sequence of the Natural 5’ Splice Site D1 on Splicing**—The intrinsic strength of a 5’ splice site may modify its usage (see “Discussion”) and to establish the rules of the relationship, we studied transcripts in which Dcr1 was suppressed by mutation of its first intronic G (Dcr1* transcripts). The sequence of D1 was then modified in such a way that its capacity to hybridize to U1 snRNA varied from 5 to 9 contiguous base pairs (Table I) and the transcripts assayed for splicing. Splicing efficiency is high as long as the number of base pairs in the potential hybrid is ≥6 (Fig. 1, B and C, see mRNA and lariat IVS for transcripts pm12, pm1, pm8, pm16, pm14) and is reduced when the hybrid is only 5 bp long (pm7). A summary of several experiments such as those illustrated in Fig. 1, B and C, and including transcripts pm18 (8 bp) and pm10 (5 bp) is presented in Fig. 1D. The data indicate that, above a certain threshold, the stability of the U1 snRNA-D1 hybrid allows full splicing efficiency (80–90%), whereas the less stable hybrids only allow splicing of a fraction of the precursor (50–60%).

The examination of the analytical gels shows that, in spite of similar splicing efficiencies, the relative amount of the D1 intermediates is higher in the transcripts harboring strong D1 sequences than in that harboring the natural D1 sequence (see, for instance, the intensity of the IVS exon 2 and exon 1 bands from pm12 and pm1 in Fig. 1C). A comparison of the amount of the D1 splicing intermediates from the various transcripts after 3 h of incubation (Table II) indeed shows a significantly higher level of intermediates in the case of transcripts able to form D1-U1 snRNA hybrids of 9 and 8 bp (pm12 and pm18) than with the other transcripts. To understand the cause of the differences, we first compared the time course of accumulation of intermediates of pm12 (9 bp), pm18 (8bp), pm8 (7 bp), pm7 (1 bp), and pm7 (5 bp) (Fig. 2B). The curves are notably different, due to a large variation of the amount of intermediates formed after the first splicing reaction (height of the peak). The variation depends, primarily though not solely (see below), on the capacity of the site to hybridize with U1 snRNA. The time lag, that is the time required for the assembly of active spliceosomes at a given splice site (Fig. 2A) is of the same order of magnitude for all.

**Table I**

*The sequences of the 5' splice sites D1 and Dcr1 from the various mutant transcripts*

<table>
<thead>
<tr>
<th>Transcripts</th>
<th>5' splice site sequences</th>
<th>Hybrids</th>
<th>D1 bp</th>
<th>Dcr1 bp</th>
<th>–kcal/mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>pm12 pm13</td>
<td>CAG:UGAU CUG:GAG 9</td>
<td>D1 Dcr1</td>
<td>11.0</td>
<td>7.6</td>
<td></td>
</tr>
<tr>
<td>pm18 pm19</td>
<td>CAG:UGAU CUG:GAG 8</td>
<td>D1 Dcr1</td>
<td>9.5</td>
<td>7.6</td>
<td></td>
</tr>
<tr>
<td>pm10</td>
<td>CAG:UGAU CUG:GAG 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm8 pm9</td>
<td>CAG:UGAU CUG:GAG 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm16 pm17</td>
<td>CAG:UGAU CUG:GAG 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm14 pm15</td>
<td>CAG:UGAU CUG:GAG 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm7 pm4</td>
<td>CAG:UGAU CUG:GAG 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm10 pm11</td>
<td>CAG:UGAU CUG:GAG 5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm3 pm2</td>
<td>CAG:UGAU CUG:GAG 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pm6</td>
<td>CAG:UGAU CUG:GAG 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table II**

*The proportion of D1 splicing intermediates in the various transcripts*

<table>
<thead>
<tr>
<th>Transcripts</th>
<th>D1 Dcr1* bp</th>
<th>-kcal/mol</th>
<th>Mean ±S.E.</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>pm12 pm13</td>
<td>CAG GUGAU</td>
<td>11.0</td>
<td>4.7 ±0.3</td>
<td>5</td>
</tr>
<tr>
<td>pm18 pm19</td>
<td>CAG GUGAC</td>
<td>9.5</td>
<td>3.6 ±0.7</td>
<td>2</td>
</tr>
<tr>
<td>pm8</td>
<td>CAG GUGAU</td>
<td>7.7</td>
<td>2.1 ±0.3</td>
<td>2</td>
</tr>
<tr>
<td>pm7</td>
<td>CAG GUGAC</td>
<td>7.5</td>
<td>1.2 ±0.3</td>
<td>2</td>
</tr>
<tr>
<td>pm16-17</td>
<td>CAG GUGAC</td>
<td>6.0</td>
<td>1.1 ±0.3</td>
<td>2</td>
</tr>
<tr>
<td>pm14-15</td>
<td>CAG GUGAC</td>
<td>4.6</td>
<td>1.8 ±0.2</td>
<td>1</td>
</tr>
<tr>
<td>pm7</td>
<td>CAG GUGAU</td>
<td>4.2</td>
<td>1.9 ±0.3</td>
<td>5</td>
</tr>
<tr>
<td>pm10</td>
<td>CAG GUGAC</td>
<td>3.1</td>
<td>1.9 ±0.4</td>
<td>3</td>
</tr>
</tbody>
</table>
transcripts (12–14 min). The kinetics of the utilization of intermediates (descending part of the curves) are similar for pm12, pm18, pm8, and pm1, and the curves may be superimposed by the appropriate translation (not shown). In contrast, for pm7, the utilization of the intermediates appears slowed. In agreement, products accumulate with similar kinetics and to similar levels for the four transcripts able to form D1-U1 snRNA hybrids of 9, 8, or 7 bp, but less efficiently with pm7 (5 bp).

From these data, we infer that the stability of the D1-U1 snRNA hybrid is a major determinant for the efficiency of the first splicing reaction. For the strongest sites, the efficiency of the second reaction appears insufficient to process the excess of intermediates formed, which explains that a high level of intermediates persists at the end of the incubation. For the weakest site (pm7), the data indicate either that the low level of intermediates is insufficient for full splicing efficiency or, alternatively, that the instability of the hybrid provokes additional changes, which perturb the progression of the second reaction.

The comparison of the time course curves of pm8 and pm1 (7 bp, Fig. 2B) and the examination of the data of Table II suggest that the intrinsic strength of the site is not the only determinant of the efficiency of the first splicing reaction. The Dcr1* transcripts are identical except for the D1 sequence and the major difference between the transcripts, besides intrinsic strength, resides in the structure of the D1-U1 snRNA hybrids, which have different 5' or 3' dangling ends. These observations suggest that the geometry of the D1-U1 snRNA hybrid may play a nonnegligible role in the activity of the splicing complexes.

Loss of the Selector Activity of the Upstream Exon upon Reduction of the Intrinsic Strength of the Natural 5' Splice Site D1—In the natural context (transcript pm0), the cryptic site Dcr1 is silent, but its presence nevertheless slightly decreases D1 usage, indicating that Dcr1 is a cis-competitor for D1 (Domenjoud et al., 1993 and Fig. 1, Dcr1* transcripts). When the strength of D1 is increased (transcripts pm13, pm19, see Table I) or is similar (pm9), D1 usage is also slightly affected, while Dcr1 remains silent. When the size of the potential U1 snRNA-D1 hybrid decreases to 6 bp (pm17, pm15) or 5 bp (pm4, pm11), Dcr1 is induced, while D1 usage decreases. The effect is small with the 6-bp hybrid but more marked with the 5-bp hybrid as about 7 and 58%, respectively, of the precursor molecules are spliced at Dcr1 in these cases. Thus, although the exonic element, shown to have a selector activity (Domenjoud et al., 1993), is present in all cases, the capacity to select D1 and to silence Dcr1, observed when the D1 sequences are able to form hybrids of 9, 8, or 7 bp with U1 snRNA, is completely lost when the potential hybrid is only 5 bp long and partially lost when it is 6 bp long.

Selection due to the presence of the exonic sequence occurs during early spliceosome assembly (Domenjoud et al., 1993). This is also the case of selection due to changes of the D1 sequence as shown by a comparative study of the time course of appearance of intermediates and products from transcripts pm0 and pm4 (7 and 5 bp, respectively, in the D1 hybrid). Very few D1 intermediates are made from pm4 (Fig. 2C, solid circles), whereas they are abundant in pm0 (Fig. 2C, squares). The comparison of pm4 and pm7, which are identical except for the absence of Dcr1 (Fig. 2, B and C, solid circles) indicates that this is the consequence of the presence of Dcr1 in pm4. Selection of D1 is evident in the 7-bp transcript (pm0) where no precursor spliced at Dcr1 can be detected, but has been lost in the 5-bp transcript (pm4) where Dcr1 is induced (triangles in Fig. 2C), with its characteristic
The amount of D1 intermediates that accumulate at the end of splicing is high in the Dcr1- transcripts with the strongest D1 sites (compare IVS exon 2 and exon 1 in pm13 and pm11, Fig. 1B), as was the case for the Dcr1+ transcripts. In fact, the Dcr1+ and Dcr1- transcripts (6 bp or more in the hybrid) behave similarly in this respect (Table II), which indicates that the presence of Dcr1 has no effect on splicing at D1, once the precursor has been committed to splicing.

The data indicate that the presence of the exonic element is not sufficient to promote the assembly of spliceosomes at D1. The sequence of D1 (which includes 6 intronic nucleotides) is also required to ensure the selector activity. The intrinsic strength of D1 has to be large enough to allow the formation of a stable hybrid (≥6 bp) with U1 snRNA. It may be recalled that the D1 sequence is not sufficient by itself for selection, since transcripts harboring the wild type D1 sequence, but lacking the exonic element, do not specifically select D1 (Domenjoud et al., 1993). Thus, the exonic element with its potential stem-loop structure (HP1) and the 5' splice site sequence cooperate for the selector activity (Fig. 3).

**The Effect of the Relative Strength of D1 and Dcr1 on Their Selection**—A high strength of Dcr1 relative to D1 could explain its dominance, as is for instance the case of transcripts pm4 and pm11, in which the selector activity is lost (5 bp in the hybrid). To determine whether the relative intrinsic strength of the sites is a parameter of selection, we prepared transcripts harboring a strong Dcr1 sequence (8 bp, −9.5 kcal/mol) and compared their splicing with that of transcripts harboring the natural Dcr1 sequence (7 bp, −7.6 kcal/mol). Transcripts in which D1 or Dcr1 were suppressed (pm1, pm2, and pm6) served as controls (Fig. 4).

We first compared pm3 (−7.5 and −9.5 kcal/mol for D1 and Dcr1, respectively) and pm4 (−4.2 and −7.6 kcal/mol for D1 and Dcr1, respectively), in which the potential Dcr1 hybrids are more stable than those of the D1 hybrids. However, the precursor is primarily spliced at D1 in pm3 and at Dcr1 in pm4 (see the lariat IVS and mRNA species in Fig. 4). This striking difference of behavior shows that selection primarily depends on the stability of the U1 snRNA-D1 hybrid (7 bp, −7.6 kcal/mol) and not on the relative strength of the two competing sites.

The importance of the relative strength of the sites for their usage can be shown when the conditions of selection are the same. If we compare pm4 and pm5, which harbor the same D1 sequence (hybrid of 5 bp, no selection of D1) but different Dcr1 sequences, we observe that D1 is almost silenced in the presence of the strongest Dcr1 sequence (pm5), whereas it is used with an efficiency of 10% in pm4 (Fig. 4). The difference of Dcr1 usage between pm4 and pm5 is similar to that observed between pm2 and pm6, that is in the absence of D1. The comparison of the time course curves shows that the overall profile of the Dcr1 curves (Fig. 2, C and E, triangles) are similar (and similar to that established previously for pm2), except for the amount of intermediates and products. Thus, when the potential D1-U1 snRNA interaction is too weak to allow selection, Dcr1 and D1 are used according to their respective strengths.

The influence of the relative strength of the two sites can also be demonstrated in the conditions of selection of D1. Transcripts pm0 and pm3 harbor the same D1 sequence (hybrid of 7 bp, allowing selection) but different Dcr1 sequences (Fig. 2, C and D). Although in pm0, D1 is entirely protected against competition by Dcr1 (7- bp hybrid), it is only partially protected in pm3 (8-bp hybrid), since about 10% of the precursor is spliced at Dcr1; moreover, splicing at D1 is lowered in pm3 relative to pm0 (approximately from 75 to 50%). In this example, the selector activity is obvious but is not sufficient to overcome the effect of the strongest cis-competitor. An enlargement of the time course of intermediates at the start of the reactions (Fig. 5) shows that the time required for spliceosome assembly gradually increases and the rate of the first reaction gradually decreases from pm1 (no Dcr1) to pm0 (Dcr1, 7 bp) and pm3 (Dcr1, 8 bp), a corollary being a reduction of splicing efficiency. As above, the relative strength of the sites influences the efficiency of splicing. Nevertheless, this effect is modest as compared with that of the cis-acting selector.

The kinetic study also indicates that D1 and Dcr1 are differently influenced by changes of intrinsic strength. For instance, the potential D1 hybrids formed from pm1 and pm18 have a stability of −7.5 and −9.5 kcal/mol, respectively. The
There are many examples showing that the sequence of a 5' splice site is important for its selection. Naturally occurring mutations of a 5' splice site (leaving the first intronic G, thus allowing the site usage) are at the origin of various genetic diseases. A survey of the data has shown that such mutations decrease the intrinsic strength of the site and allow the usage of other sites, including cryptic sites, thus inducing different splicing pathways (see for instance, Hodges and Rosenberg, 1989; Vidaud et al., 1989; Akli et al., 1990; Bonadio et al., 1990; D'Alessio et al., 1991 for mutations at G-1; Bonadio et al., 1990; Ganguly et al., 1991; Lee et al., 1991 for mutations at G+5). Similarly, a suboptimal 5' splice site sequence may be responsible for a low efficiency of splicing, and strengthening the site results in an increased usage (Talerico and Berget, 1990; Grant et al., 1990; Tacke and Goridis, 1991; Kuo et al., 1991; Hodges and Bernstein, 1992; Belaguli et al., 1992; Caffarelli et al., 1992). The importance of the ability of the site to pair with U1 snRNA has also been emphasized (Nelson and Green, 1990; Lear et al., 1990). Thus, the sequence of a 5' splice site in the context of the surrounding potential sites is a crucial parameter for selection. This is also the case for the 5' splice sites from the adenovirus E3 transcription unit that we are studying.

We have compared transcripts that are identical except for the sequence of the authentic 5' splice site D1 or the cryptic site Dcr1, so as to study the effects of changes of these 5' splice site sequences. Some of these modest changes of sequence have important repercussions on the kinetics and efficiency of splicing. Two categories of transcripts can be distinguished on the basis of our data. First, those harboring a D1 site whose capacity to hybridize to U1 snRNA is higher than that of the natural site (>6 bp). They are efficiently spliced at D1 and the cryptic site is silenced in all of them. The second category comprises transcripts with a D1 site which is only able to form a D1-U1 snRNA hybrid of 5 bp. Splicing at D1 is reduced relative to the transcripts of the first category, and the cryptic site is induced. It was shown previously (Domenjoud et al., 1993) that the selection of D1, and the silencing of Dcr1 requires an exonic element HP1, designated D1 selector. As this element is present in both categories of transcripts, our data indicate that the 5' splice site itself is part of the D1 selector. The major, if not the only, requirement of the D1 sequence for this selector activity is the ability to form a D1-U1 snRNA hybrid of more than 6 bp. The proposed secondary structure of the cis-acting selector (Fig. 3) shows that the stem-loop structure HP1 is close to the 5' splice site. It is likely that trans-acting factors recognize the selector region during spliceosome assembly. The most evident is U1 snRNP. The fact that a stable D1-U1 snRNA hybrid is indispensable for selection supports the idea that U1 snRNA and U1 snRNP are involved in the process. On the other hand, it is known that the binding of U1 snRNP is an early event in spliceosome assembly and that it is required for the formation of commitment complexes in mammalian systems (Michaud and Reed, 1991; Jamison et al., 1992). Our kinetic studies (Domenjoud et al., 1993 and present work) indicating that selection occurs during spliceosome assembly are also compatible with the involvement of U1 snRNP in the recognition of the cis-acting selector. The association of U1 snRNA to the 5' splice site will bring the U1 snRNP proteins into close proximity of HP1 (Fig. 3), and one or several proteins may bind one or several motifs of the stem-loop structure. It is plausible that these interactions are the structural basis for the observed cooperation between the D1 site and the upstream exonic element. Thus, HP1 might ensure a stability of the D1-U1 snRNP complex higher than that of a competing site and, as a consequence, D1 will be preferentially selected.

U1 snRNP contains specific proteins in addition to the Sm core proteins, common to all U snRNPs. One of these, the
U1A protein, harbors two RNA binding domains and only one of these is used for U1 snRNA recognition (Lutz-Freyermuth et al., 1990; Nagai et al., 1990; Boelens et al., 1991; Haynes, 1992); thus, the other RNA binding domain of U1A is a candidate for binding a premRNA sequence. Alternative possibilities may also be considered since various trans-acting factors have been shown to be involved in the selection of a 5' splice site. For example, the hnRNP proteins (Bennett et al., 1992), the hnRNP protein A1 and the general splicing factor ASF-SF2 (Harper and Manley, 1991; Mayeda and Krainer, 1992) or a multiprotein complex that can bind pseudo-5' splice sites (Siebel et al., 1992) participate in both early spliceosome assembly and selection. Nevertheless, the requirement for a strong 5' splice site sequence should be taken into account in any interpretation of the mechanism of the particular case of selection that we study.

Acknowledgments—We thank Dr. G. Richards for a critical reading of the manuscript. We are also grateful to the cell culture group for providing HeLa cells, to A. Staub for oligonucleotide synthesis, and to the secretarial staff for help in preparing the manuscript.

REFERENCES


Spillins, S., and Bingham, P. M. (1992) Nucleic Acids Res. 20, 5719-5727


Winship, P. R. (1989) Nucleic Acids Res. 17, 1266