Alternative Splicing of Intron 3 of the Serine/Arginine-rich Protein 9G8 Gene

IDENTIFICATION OF FLANKING EXONIC SPlicing ENHANCERS AND INVOLVEMENT OF 9G8 AS A TRANS-ActING FACTOR*

Received for publication, October 18, 2000
Published, JBC Papers in Press, November 28, 2000, DOI 10.1074/jbc.M009510200

Fabrice Lejeune, Yvon Cavaloc‡, and James Stevenin§

From the Institut de Génétique et de Biologie Moléculaire et Cellulaire, CNRS/INSERM/Université Louis Pasteur, 1 Rue Laurent Fries, BP 163, 67404 Illkirch Cedex, C.U. de Strasbourg, France

9G8 protein belongs to the conserved serine/arginine-rich (SR) protein family, whose members exhibit multiple functions in constitutive and alternative splicing. We have previously shown that 9G8 primary transcripts are subjected to alternative splicing by excision/retention of intron 3 and to a tissue specific modulation. Because both 5′- and 3′-splice sites of intron 3 appear to be suboptimal in vertebrates, we tested the 9G8 intron 3 as a novel model system of alternative splicing. By using an in vitro approach and a mutational analysis, we have identified two purine-rich exonic splicing enhancers (ESE) located in exon 4 and a (GAA)₃ enhancer located in exon 3. These elements act in concert to promote efficient splicing activation both in vitro and in vivo. Titration experiments with an excess of exonic enhancers or SR-specific RNA targets strongly suggest that SR proteins are specifically involved in the activation process. Although ASF/SF2 was expected to interact the most efficiently with ESE according to the enhancer sequences, UV cross-linking coupled or not to immunopurification demonstrates that 9G8 is highly recruited by the three ESE, followed by SC35. In contrast, ASF/SF2 only binds significantly to the (GAA)₃ motif. S100 complementation experiments with individual SR proteins demonstrate that only 9G8 is able to fully restore splicing of intron 3. These results, and the fact that the exon 3 and 4 ESE sequences are conserved in vertebrates, strongly suggest that the alternative splicing of intron 3 represents an important step in the regulation of the expression of 9G8.

Alternative splicing of nuclear pre-mRNA is a widespread mechanism for controlling gene expression in higher eukaryotes, which allows the formation of various RNA isoforms from a single primary RNA transcript (1, 2). The differential recognition and alternative choice of 5′- and 3′-splice sites, which also represent the earliest events of the spliceosome formation, may be under developmental, tissue, or sex-specific controls. These regulations are predicted to require both cis-

* This work was supported by funds from INSERM, CNRS, the Hôpital Universitaire de Strasbourg, the Association pour la Recherche contre le Cancer, and a studentship from the Ministère de l’Enseignement Supérieur et de la Recherche (to F. L.). 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: UPR 41 CNRS, Faculté de Médecine, Avenue du Pr L. Bernard, 35043 Rennes Cedex, France.

§ To whom correspondence should be addressed. Tel.: 33-3-88-65-33-61; Fax: 33-3-88-65-32-01; E-mail: stevenin@gbmc.u-strasbg.fr.

and trans-acting factors, which modulate either positively or negatively the alternative splicing process (3, 4). The study of many alternative splicing systems has allowed the identification and characterization of cis-acting elements, which activate or repress competing splicing reactions. Among them, the best characterized are the exonic splicing enhancers (ESE), which stimulate the use of nearby weak 3′-splice sites (for review, see Ref. 5). Some of these ESE are purine-rich, and they are present in various model systems. However, much less is known about the trans-acting factors and the precise mechanisms responsible for alternative splicing regulation. Important progress has been made in Drosophila models, especially for the genes involved in the sex determination pathway for which the role of several specific trans-acting factors, namely, Sex-lethal (Sxl), Transformer (Tra) and Transformer-2 (Tra-2) have also been well documented (for review, see Refs. 2, 6). However, in vertebrate systems, the advances on the identification of the cellular factors involved in splicing regulation are far more limited.

A family of essential splicing factors, the serine/arginine (SR)-rich proteins, represents one of the attractive candidates for the activation and regulation of the splice site choice in the alternative splicing process (7–9). SR proteins contain one or two N-terminal ribonucleoprotein-type RNA binding domains, which are required for RNA binding, and a C-terminal arginine/serine (RS) domain, which may promote protein-protein interactions with other components of the splicing machinery. The SR proteins are required for the constitutive splicing reaction, mainly for the formation of the early prespliceosomal complex by stabilizing U1 small nuclear ribonucleoprotein at the 5′-splice site and in spliceosome formation, most likely by mediating interactions between U1 small nuclear ribonucleoprotein and U2AF, bound to the 5′- and 3′-splice sites, respectively (7, 8, 10). SR proteins have been shown to intervene in alternative splicing both in vitro and in vivo in a concentration-dependent manner, by modulating 5′-splice site choice in pre-mRNA containing competing 5′-splice sites (11–14). They also participate in splicing activation of introns containing weak 3′-splice sites by binding to purine-rich ESE (7, 9) as well as in splicing repression (15).

To better understand the basis of their specific functions in alternative splicing regulation, analyses of the SR protein abilities to recognize specific RNA targets have been performed. High affinity RNA targets have been defined by a conventional

1 The abbreviations used are: ESE, exonic splicing enhancers; SR, serine/arginine-rich proteins; E1–E8, exons 1 through 8; Pu, purine; Py, pyrimidine; nt, nucleotide(s); kb, kilobase(s); bp, base pair(s); RT-PCR, reverse transcription-polymerase chain reaction; SELEX, systematic evolution of ligands by exponential enrichment; cons., consensus.
Splicing Activation of 9G8 Intron 3

EXPERIMENTAL PROCEDURES

In Vitro Constructions—Plasmids encoding the intron 3 pre-mRNA substrates are derived from an initial construct in which a 1118-bp SpH1/XmnI fragment of the human 9G8 gene containing the last 144 bp of exon 3, 875 bp of intron 3, exon 4, and the first 23 bp of intron 4 was cloned between the SpH1 and HinClI sites of pGEM3Zf+ plasmid (Promega). The exon 3-exon 4 (E3-E4) construct is derived from the previous wild-type minigene by generating two deletions, internal to the intron 3, which reduce its size to 469 bp. The first is a SpeI/EcoRV deletion to remove 162 bp; the second is created by an opening at the NcoI site, followed by a digestion with the exonuclease Bal31, filling-in and rejoining of the ends. In addition, an NheI site was created in the middle of exon 4 to allow us to modify easily the exon 4 sequences. The mutated exon 4 E3-E4 constructions are made by the replacement of one or both halves of the exon 4 with oligonucleotides generating the mutant exon 4-UAA, mUAA, or ΨB sequences (see Fig. 2 for the exact sequences). The constructs mutated at the exon 3 (GAA) enhancer were generated by two strategies. In the first, we created short deletions by an S1 nucleic digestion at the level of a BspMI site occurring 8 bp upstream of the exon 3/intron 3 junction. A 5-bp deletion construct (ΔI) or a 6-bp deletion was obtained (see Fig. 3A). In a second strategy, the (GAA)₆ sequence was substituted for a SpeI site by PCR-based mutagenesis, creating the ΔII construct (see Fig. 3A). Short oligonucleotides were inserted in the SpeI site to recreate one (GAA)₆ motif, in the ΔI(ΔII+GAA)₆ construct or two (GAA)₆ motifs in the ΔII+(ΔII+GAA)₆ construct. Primer sequences and further details are available on request.

The cons. 5′-E3-E4 construct contains a consensus 5′-splice site, whereas cons. 3′-E3-E4 contains consensus 5′- and 3′-splice sites. The first construct was obtained from the E3II-E4 construct in which we changed the 5′-splice site sequence AG(GTATTT to AG/GTAAGT using a PCR approach with the appropriate oligonucleotides. By the same method, we changed the 3′-splice site sequence AACTTGAAAATATG to AACTTTTTTATATG for the cons. 5′-E3-E4 construct, and the exon 4 length was decreased by a 2581 digestion prior to in vitro transcription. To obtain an enhanced exon 1′ construct, several in vitro constructs were obtained by insertion of the HindIII/EcoRI fragments from the above constructs between the HindIII and EcoRI sites of the pXJ42 plasmid (32). All clones were verified by sequencing.

To synthesize the RNA competitors corresponding to sequences of wild-type or mutated exon 4 halves, paired oligonucleotides were introduced between the KpnI/XbaI sites of the Bluescript SK vector (Stratagene). The (GAA)₆ RNA competitor, which contains the last 21 nt of exon 3 and the first G of intron 3, was obtained in the same manner. Constructs for expressing the ASF/SF2 (sequence A10: GCACAGGGAC-GAAAGACCA), SC35 (sequence 7: AGGCGCAGUAGGGUAUGC-U), and 9G8 (sequence 102: GACAACGACGACUAGAA) competitor RNA were previously described (17).

UV Cross-linking and Immunopurification Assays—For UV cross-linking, 3 μl of HeLa cell nuclear extract, S100 extract, or SR protein preparation were preincubated for 10 min at room temperature in a medium containing 50 mM KCl, 1 mM MgCl₂, 25 mM creatine phosphate, 0.76 mM ATP, 8–10% glycerol, and 100 ng of Escherichia coli tRNA, in a 10-μl volume. Interactions between proteins and RNA (400,000 cpm of [α-³²P]ATP-labeled transcript) were performed in the same medium as above, with addition of 1 mM dithiothreitol, 0.1% Nonidet P-40, 30 ng/μl bovine serum albumin, and 10 units of RNasin for 15 min at room temperature in a 15-μl volume. The reaction mixtures were then exposed to UV light (254 nm) for 10 min and subsequently treated with RNases (250 ng of RNase A, 100 units of RNase T1), for 1 h at 37 °C. Aliquots of the assays were then analyzed directly or immunopurified with monoclonal antibodies directed against ASF/SF2 or 9G8 or polyclonal antibody against SC35 as previously described (17). Each antibody recognizes its corresponding SR protein specifically. After dialution of the reactions to a volume of 100 μl with IP buffer (50 mM Tris, pH 7.5, 150 mM NaCl), the antibodies prebound to protein G-Sepharose were added and incubated overnight at 4 °C. After extensive washing, the proteins bound to the Sepharose were eluted with the 2× SDS loading buffer, heated, and then resolved by SDS-polyacrylamide gel electrophoresis on a 12% gel.

In Vitro Splicing Assays—In vitro splicing assays with 7 fmol of E3-E4 and mutated E3-E4 transcripts were performed as previously

SELEX approach for ASF/SF2, SC35, 9G8, SRp20, and SRp40 (16–18). Specific targets for SR species can contain purine-rich sequences, for ASF/SF2 for instance (16) or pyrimidine-rich sequences, for SC35 or SRp20 (17). Importantly, most of these high affinity targets act as splicing enhancers under the control of the SR species, which specifically recognizes them (16, 17). By using a complementary approach based on functional SELEX, exonic enhancers have also been identified, with more degenerate sequences that are specific of the appropriate SR protein, validating both conventional and functional SELEX approaches (19–21). Most of these prototypical motifs represent strong enhancers, which are subject to regulation according to their location in the pre-mRNA and/or their involvement in a more complex activation region. Such elements, for which interacting SR species have been unambiguously defined, include the purine-rich element of the doublesex (dsex) exon 4, recognized by ASF/SF2 (22); the Pu1 and Pu2 motifs of the bidirectional splicing enhancer of the adenoviral E1A gene, which interact with 9G8; and ASF/SF2 or SC35, respectively (23), and a constitutive splicing enhancer identified in the β-globin exon 2, specifically recognized by SC35 (24). However, as suggested recently (25), it is likely that SR proteins may also be involved in alternative splicing regulation by interacting with multiple cis-acting elements, which present only moderate affinity for one specific SR species, exhibit cross-reactivity with several SR species, or require some accessory factors to be active. Because a limited number of alternative splicing regulation examples are well established, it is important to analyze other models to get additional information.

During a previous study (26) we identified and characterized several 9G8 mRNA isoforms in human fetal tissues. Among them, the 2.4-kb isoform, containing the entire exon 3, was predominant in kidney, whereas the classic 1.4-kb mRNA isoform was predominant in liver, suggesting that a tissue-specific modulation of intron 3 splicing takes place. Further studies of intron 3 alternative splicing as a model system appeared to be interesting for several aspects. First, intron 3 of the 9G8 gene exhibits the hallmarks of an alternative intron (see also “Results”). There are only a few examples of regulated excision of introns, and the best characterized example is that of the intron 3 of the P element in Drosophila whose regulation does not involve SR species (27). Second, among the SR protein family, the transcripts of ASF/SF2 (28), SRp55, SRp40, and SRp30c (29), or SRp20 (30) have also been shown to undergo alternative splicing. Only the alternative splicing of SRp20, which occurs by inclusion/skipping of exon 4, has been analyzed in details, but the precise identification of the cis-acting elements has not yet been carried out (31). Finally, as in the case of the SRp20 gene for which the inclusion of exon 4 is expected to lead to the synthesis of a SRp20 truncated for an important part of the RS domain, the retention of the intron 3 in the 9G8 mRNA should allow the synthesis of a putative 9G8 isoform, which preserves only a reduced part of the RS domain (26). Thus, in addition to the modulation of absolute levels of the whole SR species, the alternative splicing of the 9G8 intron 3 could result in the synthesis of shorter isoforms with particular properties in splicing regulation.

Because the 5′- and 3′-splice sites of 9G8 intron 3 are intrinsically weak, we have focused our study on the cis-acting elements that control the activation of these splice sites, as well as the trans-acting factors that mediate the activation effect. We have identified strong exonic elements in exons 3 and 4, which are absolutely required for the splicing of intron 3. Moreover, we show that 9G8 is primarily involved in the splicing activation mechanism.
Splicing Activation of 9G8 Intron 3

The 9G8 Intron 3 Carries Sequences Indicative of an Alternative Intron and Its Splicing Is Modulated in Human and Mouse Tissues—Previous studies have shown that the splicing of intron 3 of the human 9G8 gene was regulated by alternative splicing in embryonic tissues (Fig. 1A and see Ref. 26). To determine whether regulation of intron 3 has been maintained during evolution, we first cloned and sequenced DNA around the exon 3/intron 3 and intron 3/exon 4 junctions in various vertebrates (Fig. 1B). As previously noted for the human 9G8 gene, the 3’-end of intron 3 of all species analyzed is particularly unusual. A purine-rich sequence of at least 10 residues is located immediately upstream to the CAG 3’, which pushes the standard U-rich sequence 15 residues upstream of the intron 3/exon 4 junction. In contrast, the 5’-splice sites appear to be less altered, although they never contain more than 5 successive consensus residues and the important G residue at position +5 is absent (in human and mouse) or flanked by two nonconsensus residues at positions +4 and +6 (Fig. 1B). Thus, this sequence comparison indicates that the 3’- and 5’-splice sites are suboptimal, in four distant species, strongly suggesting that the alternative splicing of intron 3 might occur in many vertebrates.

To test this hypothesis, we have performed Northern blot analyses with 9G8 mRNA or intron 3 probes on total RNA isolated from adult human and mouse tissues (Fig. 1, C and D). Previous analysis had only been performed on embryonic human tissues (26). As shown in Fig. 1C and as previously observed with embryonic tissues (26), four mRNA isoforms, from the classic 1.4-kb mRNA to the largest 3.6-kb mRNA, are identified with a 9G8 cDNA probe that results from alternative splicing (excision/retention) of intron 3 and use of two polyadenylation sites (see Fig. 1A for their structure). Interestingly, the alternative splicing of 9G8 mRNA is also modulated in the adult tissues. Indeed, the retention of complete intron 3, as estimated by the accumulation of the 2.4-kb isoform relative to that of the 1.4-kb mRNA, is weak in liver, lung, heart, and skeletal muscle, whereas it is highest in kidney and in pancreas (compare the left and the right panels in Fig. 1C). The same kind of splicing pattern is obtained from mouse tissues (Fig. 1D), the 2.4-/2.6-kb species doublet being poorly separated. We still observe a splicing modulation of intron 3, because the weakest retention of intron 3 was detected in skeletal muscle, whereas the relative retention of intron 3 was highest in tests, kidney, and brain. RT-PCR analysis of total RNA from various mouse tissues also revealed that, relative to the minus intron 3 mRNA isoform, the intron 3-containing isoform is predominant in kidney and brain (data not shown). Thus, data in Fig. 1 clearly show that intron 3 alternative splicing is subjected to modulation according to the tissues, i.e., modulation that could be more dramatic if considering individual cell types within tissues.
Identification of Exonic Splicing Enhancers—Because the 5′- and 3′-splice sites of intron 3 are suboptimal, we hypothesized that the splicing of intron 3 must require additional cis-acting sequences. To identify these elements, we used an in vitro splicing assay. The standard E3-E4 pre-mRNA includes the major part of exon 3, the intron 3 whose size has been reduced from 875 to 469 nt to obtain more efficient splicing (see “Experimental Procedures”) and the whole exon 4 followed by 23 nt from the intron 4. The splicing of the E3-E4 pre-mRNA was analyzed in Fig. 2B (lane 2), in which only the lariat intron 3 is shown, or in Fig. 7 (lane 2) in which both final products and free exon 3 are shown. Despite the weakness of the 5′- and 3′-splice sites of intron 3, we observed that the splicing was relatively efficient, as judged by the accumulation of the lariat intron with the remaining pre-mRNA. A similar construction (G10), which deletes one additional C nt upstream of the 5′-splice site. To analyze this possibility, progressive deletions and/or replacements of this motif were performed as well as reinsertion of various sequences (see “Experimental Procedures”). As shown above, the replacement of each purine-motif induced a decrease of the splicing reaction (Fig. 2B, lanes 6 and 7). An almost complete splicing inhibition was observed after the replacement of both purine sequences (in the mPuAB transcript, lane 8). Thus, our results suggest that the purine sequences A and B of exon 4 are directly involved and cooperate in promoting the intron 3 splicing, most likely through an activation of the weak 3′-splice site.

Immediately upstream the 5′-splice site of intron 3 (between positions −13 and −5, see Fig. 3A), we noticed a (GAA)$_3$ motif. In other alternative splicing models such motifs have been demonstrated to activate weak 3′-splice sites upstream (36, 37). In the 9G8 pre-mRNA, this motif might serve as a cis enhancer to activate the 3′-splice site of intron 2, when the intron 3 is not recognized by the splicing machinery. However, because the 5′-splice site of intron 3 is not optimal (Fig. 1B), it is possible that the (GAA)$_3$ motif is also involved in the activation of the downstream 5′-splice site. To analyze this possibility, progressive deletions and/or replacements of this motif were performed as well as reinsertion of various sequences (see “Experimental Procedures”). Typical results are shown in Fig. 3B. The construction $\Delta$I, in which the first GAAGA residues of the (GAA)$_3$ motif are deleted and only a single mutation is present in the upstream GCCGCG sequence (Fig. 3A), exhibits a strong inhibition of splicing (Fig. 3B, lanes 3 and 2, compare the accumulation of the lariat intron with the remaining pre-mRNA). A similar construction (G10), which deletes one additional C nt upstream
splicing.

common to the 3 pre-mRNA, unrelated to RNA competitor is used at half the molar A competitor, and the 9G8-specific

Fig. 2. Note that the B competitor was ing without RNA competitor. The A and B competitors correspond to the exon 4 halves A and B, respectively, as shown in Fig. 2. Note that the B competitor was used at half the molar excess as that of the A competitor, and the 9G8-specific RNA competitor is used at half the molar excess as those for ASF/SF2 and SC35. The asterisk indicates a cleavage product common to the 3 pre-mRNA, unrelated to splicing.

of the (GAA)$_3$ triplet but creates a GAGAA sequence instead of a CAGAA in the previous construct, is slightly less deleterious than ΔI (data not shown). Thus, the (GAA)$_3$ motif rather than immediately upstream sequence is primarily involved in the intron 3 activation. The substrate ΔII in which a ACUAGU replaced the (GAA)$_3$ motif (Fig. 3A) showed poor splicing as expected (Fig. 3B, lane 4). An insertion of an exogenous (GAA)$_3$ sequence resulted in a significant but incomplete reactivation (lane 5). In contrast, the insertion of two (GAA)$_3$ motifs allowed a strong splicing reactivation, because the splicing efficiency was close to that of the initial E3-E4 RNA (compare lanes 1 and 6). Together these results demonstrate that the (GAA) repeats are involved in the activation of the 5′-splice site of intron 3.

Involvement of Trans-acting Factors to Activate Splicing of Intron 3—To test whether trans-acting factors are involved in the splicing activation of intron 3, titration experiments using short RNA containing each half of exon 4 or high affinity RNA specific for individual SR proteins as competitors were performed. A control RNA in which the upstream (GAA)$_3$ motif of E3-E4 was removed and the 5′-splice site improved in a consensus site was also used (cons. 5′-E3-E4). With this RNA, the effects of competitors were expected to be linked only to the exon 4 enhancers. In a second control, the exon 4 enhancers were further deleted and the 3′-splice site was improved in a consensus site (cons. 5′3′-E3-E4). It should be noted that both control pre-mRNA (Fig. 4 lanes 12 and 20) are spliced more efficiently than the standard E3-E4 pre-mRNA (lane 2) as expected.

The results of titration experiments of the E3-E4 pre-mRNA are shown in Fig. 4. As observed with the control mPuA RNA sequence (lane 5), the half A sequence used in 750 and 1500 times excess relative to the substrate is a poor splicing competitor (lanes 3 and 4). This result was unexpected, because mutants in the half A sequence lead to inhibition of splicing (Fig. 2) and will be discussed later. In contrast, the half B sequence (lanes 6 and 7), used at twice less excess, almost completely inhibited the splicing reaction, whereas the control RNA mPuB was without effect (lane 8). Interestingly, the same behavior of the competitors was observed with the cons. 5′-E3-E4 RNA (lanes 12–16), indicating that the factors that are titrated by the half B enhancer are involved in the activation of the weak 3′-splice site. In contrast to the splicing competition revealed on these two substrates, the splicing of the cons. 5′3′-E3-E4 pre-mRNA, which lacks the identified exonic enhancers, is not significantly competed by any RNA (lanes 20–24). This demonstrates that, in our experimental conditions, the titration of the factors involved in the activation of the weak 3′-splice site doesn’t alter the multiple events of the constitutive splicing reaction analyzed with this control pre-mRNA. That one or more SR species could be involved in the activation was suggested by a specific titration assay of individual SR species. Indeed, we observed that a 9G8-specific RNA competitor identified previously (17) is as efficient as the B competitor and highly inhibits splicing of the E3-E4 and cons. 5′-E3-E4 pre-mRNA even at a moderate excess (375-fold; lanes 11 and 19). In contrast, the ASF/SF2- and SC35-specific competitors (17) were less efficient than the half B competitor (compare lanes 9 and 10 with 7 and lanes 17 and 18 with 15). As control assays, we show that the three SR-specific competitors tested individually were inefficient (lanes 25 and 26) or weak (lane 27) at inhibiting constitutive splicing of the cons. 5′3′-E3-E4 pre-mRNA. Thus, the data demonstrate that some trans-acting factors are specifically involved in splicing activation of intron 3 and suggest that the 9G8 protein itself might be one of these factors.

SR Proteins Bind to the Enhancer Sequences—To identify the putative factors that interact with the cis enhancers of exon 3 and 4, we performed UV cross-linking experiments using nuclear extracts, S100, and total SR preparations (Fig. 5A). Comparison of the cross-linking patterns with probe A reveals a specific broad band corresponding to ~35-kDa proteins, which is obtained with the nuclear extract and the SR preparation (lanes 1 and 3) but not with the S100 extract (lane 2). Significantly, the mPuA probe does not reveal such interactions (Fig. 5A, lanes 4–6). Similar results were obtained in comparing cross-linking with the B and mPuB probes (lanes 7–12), strongly suggesting that both purine-rich elements of exon 4 are effective targets for the SRp30 species of the nucleus. Finally, a comparable analysis with the (GAA)$_3$-containing probe from exon 3 shows that two specific bands correspond to proteins of ~35 and 20 kDa, indicating that some SRp30 and possibly SRp20 species interact with the upstream (GAA)$_3$ enhancer (Fig. 5A, lanes 13–15). Finally, the results presented in Fig. 5A did not reveal any significant interaction of SR species of higher molecular weight with the exon 3 or exon 4 enhancers.

Because the SRp30 proteins includes at least four individual species, we identified what species primarily bind(s) to the exon 4 enhancers. For that, an UV cross-linking with nuclear extract was followed by immunopurification analysis using specific antibodies for 9G8, SC35, or ASF/SF2 (Fig. 5B). We observed that the half A RNA interacts preferentially with 9G8.
2–4), whereas the half B RNA interacts more uniformly with the three SRp30 species but still with a preference for 9G8 (lanes 6–8). As expected, no significant interaction was revealed with the mPuA RNA (lanes 10–12). Because immunopurification experiments give qualitative rather than quantitative data, mainly due to possible differences in the affinity of the individual antibodies for their respective antigen, we developed a complementation assay in which individual SR proteins are added to an S100 extract to reconstitute splicing assays before UV cross-linking was performed (Fig. 6). The results show that 9G8 and SC35 interact efficiently with the A or B probes (Fig. 6A and B, lanes 3 and 4), the strongest interaction being between the 9G8 and B RNA probe (Fig. 6B, lane 4). In contrast, ASF/SF2 interacts moderately with B and only weakly with the A probes (Fig. 6B and A, lane 2). We have confirmed that the three SRp30 proteins interact efficiently with an RNA probe specific for each of them (Fig. 6B, lanes 10, 12, and 14) with slightly less probe cross-linked to ASF/SF2 (lane 10).

Finally, we show that the exon 3 (GAA)₃ enhancer interacts more uniformly with ASF/SF2, SC35, and 9G8 (Fig. 6A, lanes 10–12), its cross-linking pattern with the three SRp30 being similar to that of the SR-specific targets in Fig. 6B (lanes 10, 12, and 14). As expected, interactions of control mPuA or mPuB RNA are very weak with most SR species (Fig. 6A lanes 6–8, Fig. 6B lanes 6 and 7) and are only moderate for 9G8 (Fig. 6B, lane 8). Taken together, data of Fig. 6 indicate that the exon 4 enhancers interact preferentially first with 9G8 and second with SC35, whereas the exon 3 enhancer is recognized more similarly by the 3 SRp30 species. Significantly, the strongest interactions between the B half of exon 4 and 9G8 fully complement results of titration experiments (Fig. 4), showing that both the B half and 9G8-specific probes are the strongest competitors of splicing. In contrast, the weakest interaction between the A half and 9G8 might explain why the A half is a poor competitor of the intron 3 splicing.

**Activation of Intron 3 Splicing with Individual SR Species**—To analyze the capacities of individual SR species to promote the splicing of 9G8 intron 3, we carried out complementation assays containing S100 extract and the standard E3-E4 pre-mRNA or the cons. 5′-E3-E4 RNA as a suitable pre-mRNA control lacking exonic enhancers (Fig. 7). Complementation of the S100 extract with the 9G8 protein results in an efficient splicing of the E3-E4 pre-mRNA (lane 6) and of the control substrate (lane 15) as shown by the accumulations of intron 3 and mRNA relative to the unspliced pre-mRNA. In contrast to 9G8, ASF/SF2 or SC35 only poorly activates the splicing of the E3-E4 pre-mRNA (lanes 4 and 5). Unexpectedly, however, relative to the assay without exogenous SR in which a residual splicing is observed (lane 12), these two SR species do not efficiently activate the splicing of the control RNA, either individually (lanes 13 and 14) or mixed together (lane 19). Because these two SR species are active in promoting splicing of other well-characterized pre-mRNA (data not shown), the reasons why they are inappropriate in promoting
splicing of the control cons. 5’-E3-E4 remain obscure. We have confirmed that a mix of 9G8 and ASF/SF2 activates the splicing of the E3-E4 or control pre-mRNA as the 9G8 alone (compare lanes 8 with 6 and lanes 17 with 15, respectively), suggesting that ASF/SF2 and SC35 (see also lane 18) do not activate unidentified splicing silencers. However, the unexpected behavior of ASF/SF2 or SC35 species was not general. Although SRp20 does not strongly activate the splicing of the E3-E4 pre-mRNA (lanes 7 and 9), in agreement with the fact that the downstream exonic enhancers have been shown to not react with SRp20 (see Fig. 4), this SR protein is able to activate efficiently splicing of the control RNA (lanes 16 and 18). In conclusion, among the SRp30 species that interact significantly with the upstream or downstream enhancers of exons 3 and 4 of 9G8 gene, we demonstrate that only 9G8 activates efficiently the splicing of intron 3 when the SR species are tested individually. In contrast, SRp20, ASF/SF2, and SC35 appear to be unable to perform, alone or as a mix, an efficient E3-E4 splicing through an activation of its enhancers.

The Upstream and Downstream Enhancers Are Required in Vivo for Splicing Intron 3—Finally, it was important to determine whether the enhancer motifs identified in vitro are also required in vivo. For a better comparison with the in vitro experiments, we have used the same E3-E4 constructs as those used in Figs. 2 and 3, with the region of interest being inserted in a mammalian expression vector. The constructs were transfected in HeLa or NIH3T3 cells, and the splicing of the various RNA was analyzed by RT-PCR (Fig. 8). Transfection of the standard E3-E4 construct in both cells leads to the formation of the intron 3 (–) mRNA isomorph (lanes 1 and 10) that is expected for culture cells that efficiently splice the intron 3 of the endogenous 9G8 gene. Mutations within the half A result in a modest (mPuA, lane 2) or high level (YA, lane 5) of splicing inhibition. In contrast to what has been observed in vitro, mutation within the half B does not result in a detectable effect (lanes 3 and 6). However, the concomitant mutations of both halves result in a strong (mPuAB, lane 4) or in a complete inhibition (YAB, lane 7) of intron 3 splicing, similarly to what was observed in vitro (Fig. 2), indicating that the purine motifs in both halves of exon 4 have an enhancer activity in vivo and act together for the intron 3 splicing.

We have also assayed the exon 3 (GAA)$_1$ enhancer by testing ΔI (see the sequence in Fig. 3) and ΔI mPuA constructs. In HeLa cells, the (GAA)$_1$ mutation only weakly alters the splicing of the transcripts (compare lanes 8 and 9 with lanes 1 and 2, respectively). As 3T3 cells are less efficient than HeLa cells in splicing transfected 9G8 constructs (Fig. 8; compare lanes 12 with 2 for the splicing of the mPuA construct), we have done the same analysis with 3T3 cells (right panel, lanes 10-15). Interestingly, the mutation of the (GAA)$_1$ enhancer alone (lane 11) or in conjunction with mPuA or mPuB (in lanes 13 and 15) significantly impairs the splicing of the initial constructs (lanes 10, 12, and 14, respectively), indicating that the upstream enhancer has a significant role in the activation of the 5’-splice site of intron 3 and in the alternative splicing of intron 3. Thus, the data of Fig. 8 demonstrate that the identified enhancers of exon 3 and 4 also exhibit splicing activities in vivo and that they cooperate to lead to an efficient splicing of intron 3.

DISCUSSION

As observed for several pre-mRNA models, relative concentrations of SR proteins at the sites of which splicing occurs are crucial for the choice between alternative splicing reactions. As the local concentrations of SR proteins at the multiple splicing sites in the nucleus depend on their absolute levels in the cell, it is important to understand how the cell regulates the general level of each SR protein in the nucleus. Several features have prompted us to perform in this first study an identification of the cis-acting elements and trans-acting factors involved in the splicing activation of intron 3 of 9G8. First, we observed that the intron 3 splicing is modulated according to the tissues (Fig. 1), which may result in a regulation of 9G8 protein expression. Second, according to the sequences of the 5’- and 3’-splice sites of intron 3 for various vertebrates (Fig. 1), it is expected that alternative splicing of intron 3 occurs ubiquitously in vertebrates, strongly suggesting that this event is of
primary importance. Third, the 9G8 SR protein itself appears to be involved in very different splicing events. It has been shown to be specifically involved in a transcriptionally coupled alternative splicing of the fibronectin EDI exon (38), in the activation of a bidirectional splicing enhancer present in adenoviral E1A pre-mRNA (23), and in the activation of Drosophila doublesex splicing enhancer in conjunction with Drosophila Tra and Tra-2 proteins in a heterologous system (22).

Among the multiple features associated with the alternative splicing of the 9G8 intron 3 model, some appear to be unusual and other more classic. First, an alternative splicing based on an excision/retention of an intron is rather infrequent, most probably because intron retention may create a conflict between the splicing and the transport machineries. Several observations document this aspect: (i) an alternative mRNA isoform with an intron has a tendency to transiently accumulate in the nucleus before it is exported toward the cytoplasm, as shown for the tumor necrosis factor β mRNA with intron 3 (39); (ii) the efficient transport of nonspliced or incompletely spliced RNA of human immunodeficiency virus requires the involvement of the viral protein Rev, to relieve their nuclear sequestration (Ref. 40 and references therein). In this respect, it is not known whether the 9G8 mRNA isoform, including intron 3, which accumulates significantly in certain tissues (Fig. 1), is preferentially located in the cytoplasm or nucleus. In addition, the exact status of this specific 9G8 mRNA isoform, which contains a premature stop codon in the retained intron 3, relative to the mRNA surveillance mechanism (41), is not clearly defined.

Given the weakness of the 5′-splice site and mainly of the 3′-splice site of the intron 3, the identification of multiple exonic enhancers was not unexpected. In vitro and in vivo analyses show that both the exon 4 purine motifs and the exon 3 (GAA)₃ motif bear an enhancer activity for the intron 3 splicing. In in vitro assays, the presence of the three motifs is required to obtain a maximal splicing efficiency (Figs. 2 and 3). However, in culture cells (HeLa and 3T3), which perform efficient splicing of endogenous 9G8 intron 3 (data not shown), we observed that the presence of the (GAA)₃ motif may be dispensable if both exon 4 motifs are preserved, whereas the exon 4 motif B may be also dispensable in the presence of the (GAA)₃ motif and the exon 4 motif A (Fig. 8). In cells in which the splicing machinery is less optimal for the 9G8 pre-mRNA splicing, it is likely that the involvement of the three enhancers is important in promoting a significant but incomplete splicing of intron 3.

Identification of exonic enhancers similar to the (GAA)₃ motif of exon 3 that lead to an activation of a downstream 5′-splice site has been demonstrated very recently (23, 42, 43). The novel motif we identified is separated from the 5′-splice site by 4 nt (Fig. 3). In the bidirectional splicing enhancer that we previously characterized in the E1A pre-mRNA (23), only the Pu1 motif, distant by 10 nt from the 12 S 5′-splice site, exhibits an enhancer activity for this splice site. In contrast, the Pu2 motif, separated from this 5′-splice site by only 1 nt, exhibits a slight repressor activity. Therefore, from both model systems we have studied, we suggest that an upstream cis enhancer element should not be in close contact with the 5′-splice site to be fully active but has to be separated by at least 4 nt from the exon/intron limit.

In contrast, the organization of the 9G8 exon 4, with purine motifs in each half is more classic. It is reminiscent of what exists in the cassette exon 5 of troponin C, which contains in the 5′-half a first purine-rich motif important for the exon 5 recognition (37) and in the 3′-half a second purine motif, which has the capacity to dictate splice site selectivity when two 5′-splice sites are in competition (44). Several shorter purine motifs with enhancer activity are also distributed in exon 2 of α-tropomyosin (43) or in exons 6 and 8 of rat β-tropomyosin (45). We have shown that the A enhancer, although interacting less strongly than the B enhancer with SR proteins (Figs. 5 and 6) and being a weaker splicing competitor than B (Fig. 4), appears to be more efficient than B within the E3-E4 pre-mRNA at promoting splicing activation in vitro and in vivo (Figs. 2 and 8, respectively). This may be explained by the proximity of the A enhancer relative to the weak 3′-splice site of intron 3, because it has been demonstrated that the activity of splicing enhancers increases when they approach the 3′-splice site (46). Regarding the ESE we have identified in exon 3 and 4 for splicing of intron 3, it should be stressed that they could also be involved in the splicing of intron 2 and intron 4 in vivo because the retention of 875 nt of intron 3 creates an internal composite exon of 1125 nt with the flanking exons 3 and 4. Thus, it is possible that the ESE are also involved in the definition of this very large composite exon, by reinforcing its flanking 3′- and 5′-splice sites and facilitating their cross-talk.

The presence of repressor elements or splicing silencer elements has been reported in several pre-mRNA undergoing alternative splicing (15, 47–51), with some of them recruiting SR proteins (15, 51). These elements, which may be located in the vicinity of the enhancer elements, may act in a coordinated network for the on/off splicing regulation. In our case, several features reveal that the occurrence of cis repressors represents only a very remote hypothesis. First, the complete change of exon 4 by sequences of 9G8 exon 2 leads to a complete absence of splicing (Fig. 2), which is better explained by a removal of enhancer elements of the exon 4 rather than by a removal of a combination of enhancers and repressors. Second, in vivo analysis (Fig. 8), the precise mutation of only one enhancer element, A or B in exon 4 or the (GAA)₃ motif in exon 3, does not result in a significant inhibition of splicing, which could be expected if active repressors are still present within the pre-mRNA. Third, the analysis of the control pre-mRNA, in which almost all the exon 4 and the (GAA)₃ motif have been removed but the 5′- and 3′-splice sites were made consensus (the cons. 5′-3′-E3-E4), shows a very efficient splicing (Figs. 4 and 7), suggesting that no strong repressor cis activity remains in exon 3 and intron 3 sequences.

Given their richness in GA residues, the enhancers A and B of exon 4 were anticipated to be recognized primarily by ASF/SF2. Indeed, parts of enhancer A have a good match (6 of 8 nt) with the octamer AGAGAAC consensus of Tacke and Manley (16) and with the “functional” consensus (6 of 7 nt) determined by Liu et al. (19). In the same way, parts of enhancer B are a good match with the decamer consensus (16) (7 of 10 nt) or with the “functional” consensus (7 of 7 nt). The same feature was also anticipated for the (GAA)₃ motif located in exon 3. Unexpectedly, we show that ASF/SF2 interacts only moderately with the exon 4 enhancers (Figs. 5 and 6), although it recognized more efficiently the (GAA)₃ enhancer.

In contrast to ASF/SF2, 9G8 interacts efficiently with the three enhancers (Figs. 5 and 6). Interaction of 9G8 with motif A could be explained at least partly by the recognition of its 3′-part (GAAGGCAGAC), which is reminiscent of a repetition of GAC triplets as well as interaction with motif B by the presence of a GAC triplet. However, interaction of 9G8 with the (GAA)₃ motif (Fig. 6), although also resembling a (GAC)₃ motif, is not fully explained on the basis of its known target consensus (17). Interestingly, it has been shown also that 9G8 and ASF/SF2 both activate splicing of the fibronectin EDI exon, through action of an ESE, of sequence GAAGAGAGAC (38). It is possible that 9G8 protein bears a more extended capacity to interact...
with RNA targets containing G/A-rich sequences or that 9G8 is helped by coactivators that could allow its RNA binding capacities to be extended. Other lines of evidence indicate that 9G8, through interactions with exon 3 and 4 enhancers, is the primary SR species involved in the intron 3 activation. First, titration experiments by RNA targets specific for each SRp30 species demonstrate that the 9G8-specific target is the most efficient in inhibiting intron 3 splicing (Fig. 4). Second, complementation experiments with individual SR proteins (Fig. 7) show that only individual 9G8 protein promotes an effective splicing of intron 3. Although SC35 interacts significantly with the exon 3/exon 4 enhancers, this SR protein is unable to activate intron 3 splicing if added in complementation assays (Fig. 7), supporting the notion that intron 3 splicing absolutely requires 9G8 to be activated in human cells. Examination of exon 3 and 4 sequences in other vertebrates species shows that the putative enhancers have very similar sequences (data not shown). In particular, the exon 3 sequences upstream of the 5′-splice site, although encompassing a repetition of arginine residues encoded by CGX (X is any residue) or AGPu triplets, may be arranged as repetitions of GAA or GAG triplets, arguing that the activation mechanism for intron 3 splicing may be preserved in all vertebrates.

A possible mechanism for controlling the expression of splicing factors is a feedback autoregulation that may take place at the level of alternative splicing. In the case of SRp20 in mammals, the autoregulation might be negative, because its overexpression leads to the inclusion of the cassette exon 4, which contains a premature stop codon and precludes the synthesis of the classic SRp20 protein (31). In the case of the 9G8 mRNA, the retention of intron 3 should also preclude the synthesis of the entire protein (26). However, it appears that the identified exonic enhancers, which respond positively only to the 9G8 protein itself are not sufficient to promote an entire autoregulation mechanism, because the presence of the highest amounts of 9G8 protein should theoretically accentuate the splicing of intron 3 and cannot lead to a decrease of the 9G8 synthesis. In the conditions of our in vitro analysis, we were also unable to detect any antagonistic effects between 9G8 and other SR proteins (Fig. 7) such as those demonstrated previously with ASF/ SF2 and SC35 for the chicken cot- 


tehin 3 constructs will be required for a broader understanding of the alternative events and regulations occurring on the intron 3 region.

Acknowledgments—We thank Drs. R. Gattoni, J. Marie, C. F. Bourgeois, and P. Blader for fruitful discussions and critical reading of the manuscript. We are grateful to L. Kister and G. Hildwein for excellent technical assistance and the Institut de Genétique et de Biologie Moléculaire et Cellulaire core services for providing materials and technical help. We thank also Drs. P. Crigui-Filipe and S. M. Maira for cell culture advice.

REFERENCES

1. Adams, M. D., Rudner, D. Z., and Rio, D. C. (1996) Curr. Opin. Cell Biol. 8, 331–339
2. Lopez, A. J. (1998) Annu. Rev. Genet. 32, 279–305
3. Chabot, B. (1996) Trends Genet. 12, 472–478
4. Smith, C. W., and Valcarcel, J. (2000) Trends Biochem. Sci. 25, 381–388
5. Wang, J., and Manley, J. L. (1997) Curr. Opin. Genet. Dev. 7, 205–211
6. Makino, M. (1992) Curr. Opin. Genet. Dev. 2, 299–303
7. Fu, X. D. (1995) RNA (N. Y.) 1, 663–680
8. Manley, J. L., and Tacke, R. (1996) Genes Dev. 10, 1569–1579
9. Valcarcel, J., and Green, M. R. (1995) EMBO J. 14, 3540–3551
10. Bourgeois, C. F., Kister, L., and Stevenin, J. (1999) RNA (N. Y.) 5, 468–483
11. Tacke, R., Chen, Y., and Manley, J. L. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 1148–1153
12. Fu, X. D., Mayeda, A., Maniatis, T., and Krainer, A. R. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 11224–11228
13. Capecio, B. T., and Koonin, E. V. (1999) Science 285, 1783–1786
14. Bourgeois, C. F., Popielarz, M., Hildwein, G., and Stevenin, J. (1999) Mol. Cell. Biol. 19, 7547–7566
15. Schaal, T. D., and Maniatis, T. (1997) Mol. Cell. Biol. 17, 2917–2937
16. Lynch, K. W., and Maniatis, T. (1996) J. Biol. Chem. 271, 1347–1354
17. Reed, R. (2000) Annu. Rev. Genet. 34, 450–458
18. Cramer, P., Caceres, J. F., Cazalla, D., Kadener, S., Muro, A. F., Baralle, F. E., and Kornblihtt, A. R. (1999) Mol. Cell. Biol. 19, 1523–1536
19. Konig, H., Ponta, H., and Herrlich, P. (1998) EMBO J. 17, 4507–4517
20. Tanaka, K., Watakabe, A., and Shimura, Y. (1994) J. Biol. Chem. 269, 3436–3449
21. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Cell 32, 83–94
22. Kuo, M. H., and Maniatis, T. (1996) Mol. Cell. Biol. 16, 7681–7693
23. Bourgeois, C. F., Popielarz, M., Hildwein, G., and Stevenin, J. (1999) Mol. Cell. Biol. 19, 7547–7566
24. Rovei, C., Saurin, A., and Guenet, J. L. (1997) J. Biol. Chem. 272, 279–305
25. Xu, H., and Green, M. R. (1999) EMBO J. 18, 4343–4352
26. von der Houwen van Oordt, W., Diaz-Meco, M. T., Lozano, J., Krainer, A. R., Moscat, J., and Capecio, B. T. (2000) J. Cell Biol. 149, 307–316

2 F. Lejeune and J. Stevenin, unpublished results.