MOLECULAR AND CELLULAR BIOLOGY, Jan. 1993, p. 289–300
0270-7306/93/010289-12$02.00/0

Vol. 13, No. 1

Direct Selection for Mutations Affecting Specific Splice Sites in a Hamster Dihydrofolate Reductase Minigene

I-TSUEN CHEN AND LAWRENCE A. CHASIN*

Department of Biological Sciences, Columbia University, New York, New York 10027

Received 5 August 1992/Accepted 12 October 1992

A Chinese hamster cell line containing an extra exon 2 (50 bp) inserted into a single intron of a dihydrofolate reductase (*dhfr*) minigene was constructed. The extra exon 2 was efficiently spliced into the RNA, resulting in an mRNA that is incapable of coding for the DHFR enzyme. Mutations that decreased splicing of this extra exon 2 caused it to be skipped and so produced normal *dhfr* mRNA. In contrast to the parental cell line, the splicing mutants display a DHFR-positive growth phenotype. Splicing mutants were isolated from this cell line after treatment with four different mutagens (racemic benzo[c]phenanthrene diol epoxide, ethyl methanesulfonate, ethyl nitrosourea, and UV irradiation). By polymerase chain reaction amplification and direct DNA sequencing, we determined the base changes in 66 mutants. Each of the mutagens generated highly specific base changes. All mutations were single-base substitutions and comprised 24 different changes distributed over 16 positions. Most of the mutations were within the consensus sequences at the exon 2 splice donor, acceptor, and branch sites. The RNA splicing patterns in the mutants were analyzed by quantitative reverse transcription-polymerase chain reaction. The recruitment of cryptic sites was rarely seen; simple exon skipping was the predominant mutant phenotype. The wide variety of mutations that produced exon skipping suggests that this phenotype is the typical consequence of splice site damage and supports the exon definition model of splice site selection. A few mutations were located outside the consensus sequences, in the exon or between the branch point and the polypyrimidine tract, identifying additional positions that play a role in splice site definition. That most of these 66 mutations fell within consensus sequences in this near-saturation mutagenesis suggests that splicing signals beyond the consensus may consist of robust RNA structures.

The development of cell-free splicing systems has led to rapid and continuing progress in understanding biochemical mechanisms involved in intron removal and in the definition of the spliceosomal machinery. However, the exact requirements for recognition of splices sites and the rules governing splice site selection remain unclear. Analysis of consensus sequences provided the first definition of splice sites (49). Most efforts to move beyond that initial step have focused on site-directed mutagenesis studies (for reviews, see references 4 and 31). In these experiments, the effects of splice site mutations are typically assayed by using cell-free splicing extracts (in vitro) or after transient transfection of cultured mammalian cells or injection of RNA directly into frog oocytes (in vivo).

Conflicting results between these in vitro and in vivo assays for the effect of mutations on splicing have frequently been reported (2, 38, 39, 45, 83), raising the possibility that artifacts are being introduced in one system or the other. Moreover, recent experiments suggesting that the splicing machinery may be compartmentalized within the nucleus (27, 35, 41) raise the question of whether the splicing of transcripts produced by transfected DNA or of injected RNA necessarily reflects the behavior of transcripts generated in situ. For example, the use of cryptic splice sites in a human β-globin mutant is different in reticulocytes and transfected HeLa cells (44), and different cell lines sometimes produce different splicing patterns (28). Moreover, many past studies have utilized single-intron constructs, not reflecting the *cis* competition effects that are likely to influence splicing in natural multiple-intron transcripts.

An alternative, forward genetic strategy for the mutational analysis of splicing is to isolate mutants at the cell or organismic level that are defective in splicing for a particular gene. The splicing phenotype for these transcripts can then be analyzed in situ. An early example of this approach is the collection of splicing mutants found in screening human β-thalassemic individuals; the β-globin-deficient phenotypes found included many at the level of splicing (for a review, see reference 55). More recently, the induction and systematic screening of negative mutants at the dihydrofolate reductase (*dhfr*) locus in Chinese hamster ovary (CHO) cells (11, 47) and at the hypoxanthine phosphoribosyltransferase locus in a human fibroblast cell line (70, 86) have yielded dozens of mutants in which a splicing defect has been documented. While this approach is providing some new insights in RNA processing (77), it is rather laborious, as many mutants must be screened, and the splicing mutations are distributed among almost all of the splice sites in a gene. Thus, the number of deleterious mutations that reveal important *cis*-acting nucleotide positions at a particular splice site has been limited (at most four at a given splice site in the four studies cited above).

To overcome some of these difficulties, we have adopted a strategy that combines elements of genetic engineering and forward mutational analysis. We constructed a *dhfr* minigene containing three exons and two introns. The inclusion of the middle exon by splicing renders the gene product (DHFR enzyme) nonfunctional. This minigene was inserted into a CHO cell from which the endogenous *dhfr* locus had been deleted. The DHFR deficiency phenotype of these cells could be reversed only by mutations that prevent splicing of the middle exon. Selection for DHFR-positive cells yielded, at high efficiency, mutants specifically affected in splicing. In this way, we isolated, sequenced, and determined the splicing phenotypes of 66 single-base substitution mutants that

* Corresponding author.

perturbed the splicing of this single exon. This set comprised 24 different mutations; most, but not all, were within the consensus sequences of the branch, the acceptor, and the donor sites surrounding the exon. Although we isolated a relatively large number of mutants, our attempt to achieve a true saturation mutagenesis may have been thwarted by the surprisingly high sequence specificity of the four mutagens used.

MATERIALS AND METHODS

Plasmid constructions. The Chinese hamster *dhfr* minigene construct pDCH1P has been described previously (18). It contains the six exons of the gene, intron 1, about 900 bp of the 5′ flank, and the first of the two major polyadenylation sites in exon 6. Plasmid pD22 was constructed by cloning the 0.8-kb *Pst*I-*Bst*EII genomic DNA fragment of pMH8 containing exon 2 and flanks (9) into the unique *Pst*I site in intron 1 of pDCH1P. After *Pst*I site ligation, the *Bst*EII and remaining *Pst*I ends were blunted with 5 U of Klenow polymerase, 1 mM deoxynucleoside triphosphate, additional T4 DNA ligase, and overnight incubation at 15°C. pDCH1P10 was constructed by Y. Bai in this laboratory. It contains the intronless *dhfr* minigene of pDCH0 (79) cloned downstream of the T7 promoter in the riboprobe vector pSP72 (Pharmacia). A version of this minigene containing tandem copies of exon 2 was constructed as follows. Total RNA from cell line NB6 (see text) was reverse transcribed, and the full coding sequence of the *dhfr* cDNA was polymerase chain reaction (PCR) amplified as described previously (12). The 244-bp *Bal*I-*Bst*XI fragment from the PCR product was then cloned into pDCH1P10 between the *Bal*I and *Bst*XI sites. The resulting pE2 plasmid contains *dhfr* exons in the order 1-2-2-3-4-5-6.

Transfection. DNA from pD22 was linearized with *Bgl*I, and 15 to 20 ng was cotransfected with a twofold excess of plasmid pNEO-BPV100 (50) into cells of the CHO *dhfr* deletion mutant DG44 (78) by using the calcium phosphate method (84). Transfectants were selected for resistance to G418 at a final concentration of 400 μg of active compound per ml. After 9 days, colonies were cloned, expanded, and screened for pD22 insertion by PCR (see below).

DNA analysis. G418-resistant transfectant clones were screened for the presence of the cotransfecting pD22 plasmid by a rapid amplification procedure. Briefly, about 10^5 cells were centrifuged twice in 0.5 ml of water in a microfuge tube for 10 min. The drained pellet was then suspended in a standard PCR mixture containing primers from intron 1 and intron 2. PCR was performed in a thermal cycler with *Taq* DNA polymerase (Perkin-Elmer).

Ten micrograms of genomic DNA from PCR-positive clones and from CHO UA21 cells (containing one copy of the endogenous *dhfr* gene) was digested with *Eco*RI and analyzed by Southern blotting (69). A mixture of two probes was used for hybridization: a 1.4-kb *Eco*RI fragment from pD22 that spans the 5′ end of the *dhfr* gene up to exon 2 and a 2.7-kb *Eco*RI fragment that spans the 3′ half of the hamster gene for adenine phosphoribosyltransferase (*aprt*) in plasmid pH2 (46). The latter served as an internal control for determining gene copy number. Probes were labeled with ^{32}P by using a random priming kit (Boehringer Mannheim). The diagnostic autoradiographic bands were quantified by densitometry.

Mutant isolation. The parental cell line used for mutagenesis was NB6, a cell line containing a single copy of the pD22 construct and exhibiting a DHFR deficiency growth phenotype, an inability to grow in medium lacking the end products of one-carbon metabolism. For the isolation of spontaneous mutants, 24 separate populations of NB6 cells were established from inocula of 50 cells to ensure the independence of subsequently isolated mutants. When the cultures approached confluence, approximately 5×10^5 cells from each culture were challenged in purine-free medium in 100-mm dishes to select for DHFR-positive revertants, essentially as described previously (75). Mutagenic treatments were also performed on separate cultures (typically 15) established from small inocula. The mutagens used were (i) ethyl methanesulfonate (EMS), 2.4 mM for 18 h; (ii) racemic benzo[c]phenanthrene diol epoxide (BcPHDE; kindly provided by R. Harvey, University of Chicago), 0.15 μM for 90 min; (iii) ethyl nitrosourea (ENU), 3 mM for 60 min; (iv) irradiation for 20 s with a 30-W germicidal UV lamp 50 cm above dishes with the lid off. The dose of mutagens used was designed to produce 60 to 70% cell killing on the basis of prior toxicity tests. After 3 days of expression in complete medium, 5×10^5 survivors from each culture were challenged on 100-mm dishes of purine-free medium. Colonies were isolated after 7 days; only one colony per dish was cloned, recloned, and further analyzed.

DNA sequencing. Genomic DNA of each mutant was used for PCR amplification. Two intronic primers (20-mers), one beginning 31 bases upstream and the other 136 bases downstream of exon 2A of pD22, were used in most cases. For mutants affected at the branch site, a 5′ primer further upstream was used, located 57 bases upstream of the translational start (see reference 48 for sequence). Direct DNA sequencing of the PCR-amplified templates with *Taq* polymerase (Promega) and ^{32}P-end-labeled primers was carried out as previously described (12); the primer used was a 20-mer from the transcribed strand starting 68 bases downstream of exon 2A. For some mutants whose base changes were ambiguous, a direct sequencing method with Sequenase was used (63), with some modifications (15). Briefly, after PCR products were purified with Geneclean (BIO 101), 10 pmol of sequencing primer was added and the mixture was boiled for 5 min, and then quickly cooled in an ice slurry (no annealing step). Sequencing was then carried out by using a Sequenase kit (USB) and [^{32}P]dATP incorporation.

RNA analysis. Total RNA was prepared (17) and analyzed for the inclusion and exclusion of exon 2A. For qualitative analysis, reverse transcription followed by PCR amplification (RT-PCR) of *dhfr* mRNA was carried out as described previously (12). To determine the splice site of mRNA molecules spliced at a cryptic site, fragments were isolated from an agarose gel, purified with Geneclean, and then reamplified prior to direct sequencing as previously described (12).

For quantitation of RNA by PCR, a 20-μl reaction mixture containing 1 μg of total RNA (in water), 0.4 μg of random hexamers, 10 mM dithiothreitol, 40 U of RNasin, 0.5 mM deoxynucleoside triphosphate (previous four from Promega), 4 μl of 5× RT buffer and 200 U of mouse mammary leukemia virus reverse transcriptase (above two from Bethesda Research Laboratories) was incubated at 37°C for 60 min, heated to 95°C for 5 min, and then quick chilled in ice. The cDNA mixture was used directly for PCR or stored at −70°C. Half of the cDNA was used for PCR with *dhfr* primers and half was used with *aprt* primers. To a standard 50-μl PCR mixture was added 4 μCi (0.4 μl) of [^{32}P]dATP (3,000 Ci/mmol; Amersham). PCR was performed for 16 to 24 cycles, depending on the initial amount of *dhfr* mRNA present. Three microliters of each PCR sample was electro-

phoresed in a 5% polyacrylamide gel. The gel was dried, and the radioactive bands were visualized by autoradiography and then quantified on a Phosphoimager (Molecular Dynamics).

To demonstrate that the different species of *dhfr* RNA can be measured quantitatively by RT-PCR, a reconstruction experiment using RNA synthesized in vitro was carried out. *Bam*HI-linearized pDCH1P10 and pE2 (250 ng each) were incubated with 0.5 μCi of [α-^{32}P]UTP (about 1% of the amount used to make riboprobes), 0.8 mM UTP, T7 RNA polymerase (USB), and other reagents for RNA synthesis as previously described (10). After synthesis, DNase treatment, and extraction, the RNA was precipitated with ethanol and dissolved in water. The radioactivity in each preparation was counted, and the amount of RNA was calculated. The two RNA preparations were mixed in various ratios, with the total amount kept at 50 pg. The mixtures were then amplified by RT-PCR and quantified by Phosphoimaging as described above.

RESULTS

Isolation of a transfectant that allows the direct selection of splicing mutants. Our basic strategy was to create a *dhfr* minigene containing an extra exon, the incorporation of which would disrupt the coding properties of *dhfr* mRNA (Fig. 1A). By transfection, the disrupted minigene would be placed in a DHFR-deficient mutant Chinese hamster ovary (CHO) cell deleted at the *dhfr* locus; since these cells carry a mutant minigene, they would remain DHFR deficient. After mutagenesis, DHFR-proficient colonies would be selected; these revertants should include mutations that cause the offending exon to be skipped, thus restoring an uninterrupted DHFR mRNA. Sequencing the pertinent region of PCR-amplified revertant minigenes should define positions necessary for efficient splicing. An assumption in this strategy is that most mutations that decrease the splicing efficiency of the extra exon also provoke exon skipping. Exon skipping is the predominant consequence of splicing mutations arising in situ, and in particular in the *dhfr* gene (11, 47).

We started with a minigene that retained only the first of the five original *dhfr* introns (pDCH1P [18]). Into this intron 1 we inserted an extra copy of exon 2 (exon 2A), resulting in a minigene (pD22) that has the structure E1-intron-E2A-intron-E23456 (Fig. 1B). An exon from the *dhfr* gene itself was chosen with the aim of maximizing the splicing in of the extra exon; the E1-E2A splice mimics a splicing event of the normal pre-mRNA, although the E2A-E2 involves a novel exon joining. The 50-base size of this exon ensures the eradication of enzyme activity, since the resulting frame shift would lead to early truncation of the protein. For this scheme to work, the splicing in of exon 2A must be nearly 100% efficient, as CHO cells that contain as little as a few percent of wild-type enzyme levels display a DHFR-positive growth phenotype (20, 79). As a preliminary test of the efficiency with which exon 2A was spliced in, equal amounts of pD22 and the original wild-type minigene, pDCH1P, were compared for the ability to confer purine prototrophy on DG44 cells. DG44 is a CHO mutant (78) that has had both of its *dhfr* alleles deleted; the lack of this enzyme activity results in a triple auxotrophy for glycine, thymidine, and a source of purines. Although a few colonies did appear after transfection with pD22, their frequency was about 1% that of pDCH1P (data not shown), suggesting that the inclusion of the extra exon was indeed disrupting the *dhfr* coding sequence.

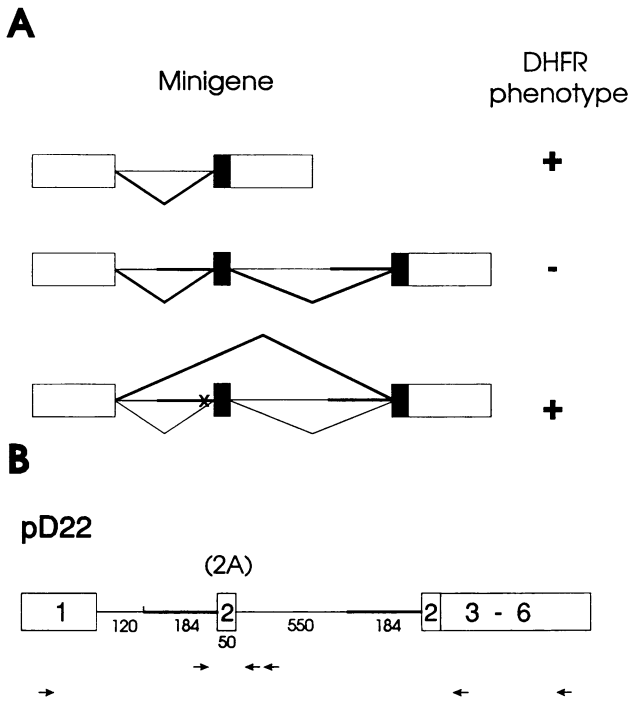

FIG. 1. Strategy for the direct selection of splicing mutants. (A) Diagram of the *dhfr* minigenes used or created, with their associated phenotypes. Boxes represent exons; the left open box is exon 1, the right open box is exons 3 to 6, the filled box is exon 2. The lines represent introns; the thicker lines indicate sequences repeated in gene construction. The angled lines illustrate splicing patterns. The figure is not drawn to scale. The top line depicts a single-intron *dhfr* minigene (pDCH1P) described previously; transfection of this or similar genes into DHFR-deficient CHO cells (DG44) confers a DHFR-positive growth phenotype, the ability to grow in purine-free medium (18, 79). The second line shows the derivative minigene constructed here (pD22). *dhfr* exon 2 together with surrounding intronic sequences were inserted into the intron of pDCH1P. The extra exon 2 (exon 2A in the text) is spliced into the mature mRNA, as shown, resulting in an mRNA that is incapable of coding for active DHFR; a CHO DG44 cell transfectant (NB6) bearing a single copy of this gene retains its DHFR-negative growth phenotype. The third line represents a resident pD22 minigene that has been mutated after a mutagenic treatment of NB6 cells. The example shows a mutation at the acceptor site upstream of exon 2A and the partial mutant splicing phenotype that can result: continued correct splicing (below) as well as the skipping of exon 2A (above). The latter generates an mRNA that is once again capable of coding for active DHFR, and so the mutant minigene now confers a DHFR-positive growth phenotype, and the mutant cell can be selected by growth in purine-free medium. (B) Map of pD22. The vertical line in intron 1 denotes the *Pst*I site into which the genomic fragment spanning exon 2 (approximately 780 bp) was introduced. Arrows show the location of primers used for PCR and sequencing; the top line shows genomic sequencing and the bottom line shows RNA amplification and sequencing. On each line the middle arrow represents the sequencing primer.

pD22 was then cotransfected with a plasmid carrying the *neo* gene into DG44 cells. G418-resistant transfectants were screened for the presence of the pD22 minigene by PCR with primers from the intron regions flanking exon 2A (Fig. 1B). Five of 35 G418-resistant cell lines contained at least this region of the pD22 construct. To ascertain the amount and arrangement of integrated pD22 sequences in transformed

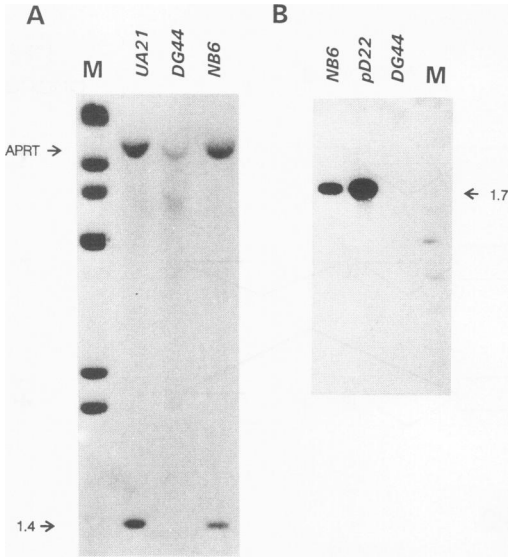

FIG. 2. pD22 sequences in transfectant NB6. (A) Southern blot showing the presence of transfected *dhfr* DNA. DNA from the indicated cell lines was digested with *Eco*RI and hybridized to a mixture of two probes, one detecting a fragment from the *aprt* gene (top band) and one that detects a 1.4-kb *Eco*RI fragment spanning the 5′ end of the hamster *dhfr* gene. The *dhfr* probe should also detect fragments from the 3′ duplicated region that includes the downstream exon 2, but these are too small to be detected in this gel. UA21 carries one copy of the endogenous *dhfr* gene, DG44 is a *dhfr* deletion mutant, and NB6 is a clone isolated after cotransfection of DG44 cells with pD22 DNA and a plasmid carrying the *neo* gene. The marker lane (M) contains end-labeled *Hin*dIII fragments of bacteriophage lambda DNA. (B) PCR amplification of transfectant DNA. Primers from *dhfr* exons 1 and 6 were used to amplify *dhfr* DNA sequences. Transfectant NB6 carries the 1.7-kb fragment characteristic of the pD22 construct. DG44 is the CHO *dhfr* deletion mutant host cell line. The marker lane contains end-labeled *Hae*III fragments of ϕX174 DNA.

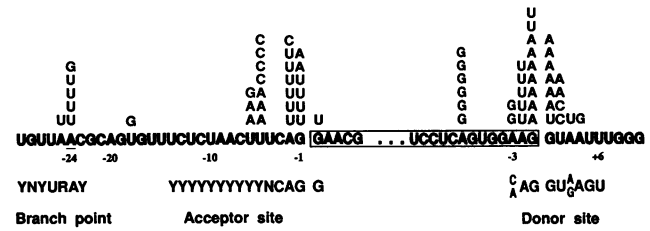

FIG. 3. Summary of base changes in mutant clones. The DNA sequence in the region surrounding exon 2A of the *dhfr* minigene is shown, with the exon boxed. Numbering is relative to the nearest splice site. The single-base substitutions found in the mutants are indicated above the wild-type sequence. The consensus sequences at or near splice sites for mammalian cells are indicated below. The mutagen used to induce each change can be found by examining the corresponding position in Fig. 4.

cells, DNA from these five lines was analyzed by Southern blotting. In the blot illustrated (Fig. 2A), a diagnostic 1.4-kb *Eco*RI fragment was absent in the DNA from the DG44 deletion mutant but was observed in DNA from transfectant NB6 and from CHO UA21 cells; the latter carry a single copy of the *dhfr* gene (UA2 [76]). Densitometric analysis showed that the ratio of the 1.4-kb *dhfr* band to a 10-kb internal control band from the adenine phosphoribosyltransferase gene was very similar in NB6 and UA21 DNA, suggesting that the NB6 cell line contains a single copy of the pD22 insert. The remaining four PCR-positive clones contained either multiple copies or a truncated insert (data not shown).

To verify the intactness of the insert in NB6 cell line, primers located in exons 1 and 6 that flank the *dhfr* coding sequence (Fig. 1B) were used to PCR amplify this region from NB6 and pD22 DNA; products of identical size were produced (Fig. 2B). In addition, in the case of NB6, the region surrounding exon 2A was sequenced from the PCR product and found to be identical to the known *dhfr* sequence (data not shown). Finally, RT-PCR with primers from exons 1 and 6 yielded a product of the expected size, a *dhfr* mRNA that was 50 bases longer than the corresponding wild-type product. However, the level of the extra 50-base mRNA observed in NB6 was only about 10% of that seen in UA21 cells (see Fig. 6, lane 1 versus lane 3). Only trace amounts of mRNA of wild-type size could be detected, indicating little or no skipping of the exon 2A.

Spontaneous DHFR-positive revertants. In accord with the RNA splicing result, NB6 cells exhibited a DHFR deficiency phenotype, failing to grow in a purine-free medium. When large numbers (10^5 to 10^6) of NB6 cells were challenged in this selective medium, spontaneous DHFR-positive revertant colonies appeared at a frequency of about 2×10^{-5}. Thirteen independent revertants were characterized by PCR amplification and sequencing of the region surrounding exon 2A. Two classes of revertants were found. Revertants of the first class (2 of 13) had deleted the duplicated portion of the pD22 minigene DNA, a result that was confirmed by Southern blot analysis (data not shown). The deletions restored the structure of the original pDCH1P minigene and presumably arose by intragenic homologous recombination. The second class consisted of point mutants. In five mutants an A-to-G base substitution had occurred at the −2 position of the splice acceptor site that preceded exon 2A, altering the highly conserved AG dinucleotide of the acceptor consensus sequence. The remaining six mutants were altered in a more surprising way; all six carried an A-to-G base substitution within exon 2A, 8 bases upstream of the splice donor site. For several of these mutants, the A-to-G base change was confirmed by sequencing the RT-PCR-amplified *dhfr* RNA.

Mutagen-induced splicing mutants. In an attempt to saturate the exon 2A splice sites with mutations affecting splicing, we treated NB6 cells with four different mutagens and selected 59 independent DHFR-positive revertants. The mutagens used, BcPHDE, EMS, ENU, and UV light, killed about 60% of the population and yielded DHFR-positive colonies at an frequency of about 3×10^{-4}, an induction of 10-fold over the spontaneous frequency. For all 59 mutants, the DNA sequence in the exon 2A region was PCR amplified from genomic DNA and directly sequenced. Two of the mutants gave no PCR product with the intronic primers used, reflecting perhaps a deletion of these intron sequences along with exon 2A or a large insertion that precluded PCR amplification. These two revertants have not been pursued further. Of the 57 induced mutants, 55 contained single-base changes and 2 were intragenic recombinants. Of the 66 point mutants isolated (including spontaneous revertants), 35 were affected at the splice donor region, 25 were in the acceptor region (including 11 in or near the polypyrimidine tract), and 6 were at or near the presumed branch point. The exact base changes obtained are shown in Fig. 3 and in Table 1. Most of the mutations occurred within the donor or acceptor consen-

TABLE 1. Characteristics of splicing mutants

Name	Mutagen	Position[a]	Base change[b]	Relative RNA level[c]	Exon 2A splicing (%)[d]	Consensus score[e]	U1 RNA binding energy[f]
NB6 (wild type)		Branch point	UGUUA*A*CG	100	100		
NS13	BcPHDE	−25	UGUU**U***A*CG	89	85		
NS25	BcPHDE	−24	UGUUU**U**CG	27	29		
NS14	BcPHDE	−24	UGUUU**G**CG	42	4		
NS76	ENU	−18	U → G	58	55		
NB6 (wild type)		Acceptor site	UUUCAG/G	100	100	84.0	
NS20	UV	−6	**A**UUCAG/G	29	48	78.2	
NS94	UV	−6	**G**UUCAG/G	26	17	78.4	
NS4	UV	−5	U**A**UCAG/G	72	8	78.2	
NS10	UV	−5	U**C**UCAG/G	130	48	82.5	
NS3	BcPHDE	−2	UUUC**C**G/G	39	0	70.2	
NS6	BcPHDE	−2	UUUC**U**G/G	70	0	70.2	
NS2	BcPHDE	−1	UUUCA**U**/G	200	0	70.2	
NS75	ENU	−1	UUUCA**A**/G	23	0	70.2	
NS28	BcPHDE	+1	UUUCAG/**U**	63	48	78.1	
NB6 (wild type)		Donor site	AAG/GUAAUU	100	100	86.3	−7.4
NS1	Spontaneous	−8	A → G	78	68	86.3	−7.4
NS79	ENU	−3	**G**AG/GUAAUU	70	0	83.3	−7.3
NS42	Spontaneous	−2	A**U**G/GUAAUU	57	0	78.0	−5.6
NS92	ENU	−1	AA**U**/GUAAUU	33	32	73.4	−2.3
NS8	EMS	−1	AA**A**/GUAAUU	26	0	73.7	−2.5
NS5	ENU	+1	AAG/**U**UAAUU	300	0	69.3	−1.4
NS7	EMS	+1	AAG/**A**UAAUU	100	0 (6[g])	69.3	−1.4
NS82	ENU	+2	AAG/G**A**AAUU	95	0 (5[g])	69.3	−3.2
NS60	BcPHDE	+2	AAG/G**C**AAUU	170	19 (4[g])	69.3	−3.2
NS91	BcPHDE	+3	AAG/GU**U**AUU	300	0	77.0	−4.9
NS90	UV	+4	AAG/GUA**G**UU	82	6 (9[g])	75.8	−6.4

[a] Acceptor −1 refers to the 3′ end of intron 1 and donor +1 refers to the 5′ end of intron 2.
[b] Mutations within consensus sequences are shown in boldface and underlined in context. The putative branch point is italicized.
[c] The sum of the radioactivity in all *DHFR* bands normalized to *APRT* and expressed relative to this value for NB6.
[d] Percentage of transcripts that include exon 2A. The limit of detection is about 2%.
[e] Calculated by the method of Shapiro and Senapathy (65) with the data of Senapathy et al. (64).
[f] Standard free energy change of duplex formation with the 9-base U1 small nuclear RNA sequence complementary to the splice donor consensus. Thermodynamic parameters were from Freier et al. (25). Calculations excluded bulges and isolated base pairs, and included G · U base pairs, loops, and terminal base pairs or mismatches. The terminus for a helix that extended to the end of the 9-base consensus was considered to be at the next base beyond this boundary (74).
[g] Use of a cryptic site at position +31 in intron 2.

sus sequences, but some fell outside this region, at −8 relative to the donor splice site and at −18 relative to the acceptor site. Some consensus bases that might have been expected to have been hit were spared, such as position +5 and +6 relative to the donor and −3 at the acceptor site. In all, 24 distinct mutational changes were found, encompassing 16 sites. Most sites were hit multiple times.

BcPHDE, EMS, ENU, and UV have been shown to induce predominantly point mutations in mammalian cells (13, 29, 51, 77, 80). Moreover, BcPHDE has been found to induce mutations preferentially at splice acceptor sites in the *dhfr* gene in CHO cells (13). As shown in Fig. 4, it is apparent that these mutagens induced point mutations with marked position dependence. Each mutagen produced multiple hits at hot spots peculiar to that mutagen. ENU preferentially attacked donor sites, and EMS did so exclusively. In contrast, UV affected mainly thymine dimers in the polypyrimidine tract, whereas BcPHDE hit both the branch point and the acceptor site. The highly nonrandom action of this set of mutagens has undermined our attempt at using saturation mutagenesis to define all sites needed for efficient splicing.

Quantitative analysis of RNA by using RT-PCR. To evaluate the mutant splicing phenotypes we quantified the amount of correctly and incorrectly spliced *dhfr* mRNA in the revertant cell lines. For this purpose we used RT-PCR methodology, with two features that enhanced the quantitativeness of this assay. First, we used [α-^{32}P]dATP in the

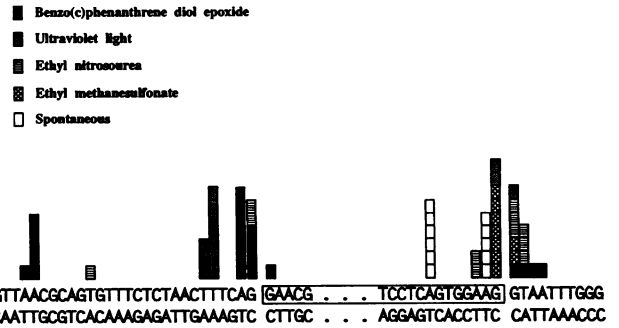

FIG. 4. Mutagen specificity of splicing mutations. Each box represents an independent mutant. The top strand represents the nontranscribed strand, with the sequence corresponding to exon 2A boxed. The base change that has occurred in each mutant can be found by examining the corresponding position in Fig. 3.

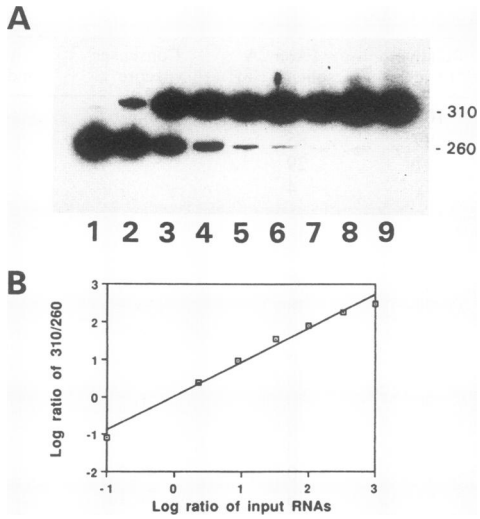

FIG. 5. Quantitation of differentially spliced products by RT-PCR. Two different plasmids (see Materials and Methods) were used as templates for the in vitro synthesis of simulated mRNA molecules that contained either one (E1) or two (E2) copies of exon 2. The two RNA molecules were mixed in various ratios to give a total of 8 pg and then subjected to reverse transcription by using random hexamers as primers. The resulting cDNA was PCR amplified for 16 cycles in the presence of ^{32}P-labeled dATP, and the PCR products were separated by native polyacrylamide gel electrophoresis. (A) Lane 1, 100% E1 (260-bp product); lanes 2 to 8, mixtures with the following ratios of E2/E1: 0.1, 2.3, 9, 32.3, 99, 332, and 999; lane 9, 100% E2 (310-bp product). (B) The radioactive bands shown in panel A were quantified by using a Molecular Dynamics Phosphoimager. The ratio of the radioactivity in the 310-bp band to the 260-bp band is plotted against the input ratio of lanes 2 to 8. The log-log plot shows that the radioactivity measurements yield a faithful rendition of the input ratios over a range of 4 orders of magnitude.

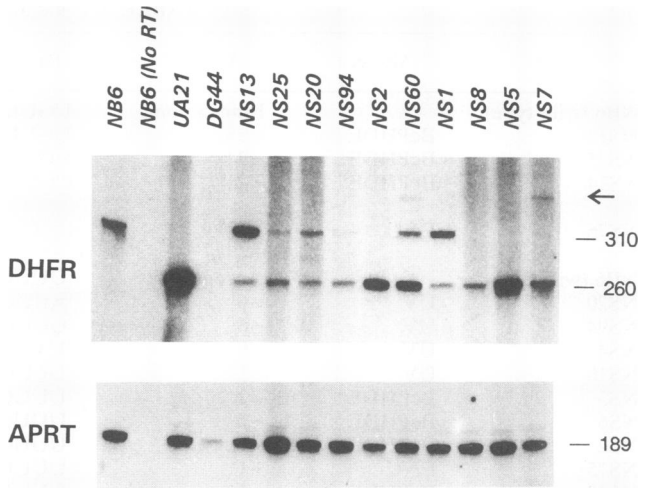

FIG. 6. *dhfr* mRNA splicing products in mutant cell lines. Total cellular RNA from the indicated cell lines was amplified by RT-PCR, the PCR product being radioactively labeled by the inclusion of [^{32}P]dATP in the PCR step. After reverse transcription, the cDNA was divided in two, with one half amplified with *dhfr* primers and the other with *aprt* primers. After separation by native polyacrylamide gel electrophoresis, the dried gels were subjected to autoradiography (shown) and phosphoimaging to generate the data shown in Table 1. The 260-bp band represents mRNA molecules that exclude the extra exon 2A, a product identical to that of authentic *dhfr* mRNA (UA21 cells). The 310-bp band represents molecules that include exon 2A, as in the original NB6 transfectant carrying a single copy of the pD22 plasmid. It can be seen that NB6 cells produce only the exon 2A-containing mRNA and that the level of this mRNA is about 10% that of the normal mRNA produced by UA21 cells, which carry a single copy of the endogenous *dhfr* gene. The *aprt* band provides a control for the reverse transcription and for the PCR reagents and was used to normalize the data in Table 1. The low level of *aprt* mRNA seen in the deletion mutant DG44 in this experiment is not typical of this cell line, and longer exposure of this autoradiogram failed to show any *dhfr* signal in DG44. The band indicated by the arrow represents the product of splicing from a cryptic donor site (see text).

PCR to produce a radioactive product that could be detected with great sensitivity. This sensitivity allowed us to stop the PCR while it was still in its exponential phase, before large amounts of product (e.g., amounts detectable by ethidium bromide binding) had accumulated. The use of the radioactive label also enabled us to use a Phosphoimager to quantify the PCR product bands after electrophoresis. Second, we amplified the same cDNA preparation with primers for adenine phosphoribosyltransferase (*aprt*) mRNA, which served as an internal control for the reverse transcription and a parallel control for the PCR.

To assess the quantitative nature of this method, we constructed recombinant riboprobe plasmids that incorporated either wild-type *dhfr* cDNA or a cDNA with the extra exon 2A. After generating each of these RNA molecules by in vitro transcription, they were mixed in widely varying proportions and subjected to the RT-PCR procedure, with primers from exon 1 and exon 4. An autoradiogram of the results is shown in Fig. 5A. The two PCR product bands of 310 and 260 bp differ by the size of the extra exon 2A. The autoradiogram was overexposed in order to visualize weak bands. In Fig. 5B is shown a plot of the ratio of the Phosphoimager values of the larger to the smaller band, representing the splicing in of exon 2A and the skipping of exon 2A, respectively. A log-log scale was used to display the wide range of ratios tested. The results indicate that the RT-PCR procedure faithfully reproduces the relative inputs of RNA, that the presence of the shorter template did not interfere with the amplification of the longer template, and that exon skipping in less than 1% of these mRNA molecules can be detected.

Quantitation of spliced RNA in the revertants. This RT-PCR assay was used to detect and quantify the splicing patterns exhibited by these revertants. An autoradiogram representative of these results is presented in Fig. 6. As expected, no *dhfr* band was observed in the *dhfr* double-deletion mutant DG44 (lane 4). The parental NB6 cell line containing a single copy of the pD22 plasmid expressed the 310-bp PCR product characteristic of *dhfr* mRNA that includes exon 2A. The level of this RNA was less than 10% that seen in the UA21 cell line, which carries one copy of the endogenous *dhfr* gene. This relatively low level of expression is probably unrelated to the unusual exon arrangement in pD22, since similarly low RNA levels in transfectants have been observed by using more orthodox *dhfr* minigenes (18, 77, 79) and may be caused by a position effect or the lack of a full set of introns. A revertant capable of growth in purine-free medium must contain wild-type *dhfr* mRNA, i.e., one that has excluded exon 2A and which produces a 260-bp product in this PCR analysis. In mutants such as NS2, 8, and 5 in Fig. 6, the 310-bp band was not detectable, having been completely replaced by the 260-bp band, indi-

cating less than 2% exon 2A splicing relative to the parental NB6 cells. Leaky mutant phenotypes were also observed, i.e., revertants that produced both versions of mRNA, such as mutants NS13, 25, 20, 94, and 1 in Fig. 6. Less common was a third phenotype, in which an additional minor band of about 340 bp was observed (NS60 and 7 in Fig. 6). This molecule is the result of splicing at a cryptic donor site 31 bases downstream of exon 2A (see below).

The radioactive RT-PCR products were quantified by Phosphoimaging, normalizing to the level of *aprt* mRNA in each preparation. The results for 24 revertants, representing each of the different mutational changes observed, are presented in Table 1. Table 1 presents the level of total *dhfr* mRNA in each revertant (up to three species summed) as well as the proportion of transcripts that have persisted in the splicing in of exon 2A; in both cases the values are expressed relative to the parental NB6 cell line.

At the acceptor site, all four mutations at the universally conserved positions, −1 and −2 (NS2, 3, 6, and 75), resulted in complete exon 2A skipping (no detectable exon 2A inclusion). A change at the less-well-conserved +1 position was more permissive, with about equal amounts of inclusion and skipping (NS28). Base changes found in the proximal region of the polypyrimidine tract (positions −5 and −6) retained a variable amount (8 to 48%) of exon 2A splicing (NS4, 10, 20, 94). Notable among these is NS10, which while retaining a pyrimidine at −5 nevertheless skips exon 2A half the time. From the high number of hits at position −24 and the agreement of the surrounding sequence with the branch point consensus (YNYURAY [53]), it is likely that the A at this position represents the branch point for exon 2 and 2A. Base changes at this branch point (NS14 and 25) reduced exon 2A splicing by as much as 96%. At the position 1 base upstream from the branch point (−25), an A-to-U change had a small but discernable effect (NS13). The 15% exon 2A skipping exhibited by this mutant demonstrates that even a relatively minor splicing deficiency can be selected for in this system. Finally, a U-to-G substitution at position −18, beyond the usually accepted limits of the polypyrimidine tract (−5 to −14 [64]) and the branch point consensus (−29 to −24 here) reduces splicing by about half.

At the donor site, 3 of 4 mutations at the well-conserved GU dinucleotide region (positions +1 and +2) obliterated correct splicing (Table 1, NS5, 7, and 82). The fourth mutant (NS60), with a change from GU to GC, continued to splice exon 2A into 20% of its transcripts. In addition, 3 of these mutants (NS7, 60, and 82) activated a cryptic splice site, AGG/GUUUGU, located 31 bases downstream of the normal donor splice site, resulting in an mRNA with a longer exon 2A than the parental line. The identity of this cryptic site was ascertained by gel purification of the resulting high-molecular-weight PCR band followed by reamplification and direct sequencing (data not shown). This cryptic site was also utilized by a mutant affected at position +4 (NS90). On the other hand, mutations at 7 other positions at or near the donor site that produced exon skipping did not provoke the use of the +31 site (Table 1).

Although the G and U at positions +1 and +2 are nearly universally conserved, the remaining positions of the consensus are less so, ranging from 47 to 94% in rodents (64). Mutations at 5 of the 7 less-well-conserved positions were found; more often than not, the phenotype was severe, with little or no correctly spliced transcripts detected (NS79, 42, 8, 91, and 90 at −3, −2, −1, +3, and +4, respectively). No mutants were isolated that affected positions +5 or +6. Finally, one position near the donor site fell well outside the 9-base consensus. An A-to-G change at position −8 within the exon failed to splice exon 2A 32% of the time.

It can also be seen in Table 1 that the steady-state level of *dhfr* mRNA species varied over a 12-fold range among the mutants, from 25 to 300% of the parental NB6 level. It is not clear what the determining factors are that influence this final level, as low and high expressors are found among both donor and acceptor mutants, and both null or leaky splicers. Mutations in the donor intron sequences at positions +1 to +4 always produced a normal or higher level of RNA.

DISCUSSION

In vivo selection of splicing mutants. Having created the NB6 cell line harboring a *dhfr* minigene with a disrupting exon, we were able to readily select mutants in which the splicing of the offending exon was diminished. The sensitivity of this selection is illustrated by mutant NS13, which preserves correct splicing of pre-mRNA at 85% of the parental level. Exon skipping in this mutant results in functional *dhfr* mRNA that is only 13% of the amount of nonfunctional parental NB6 mRNA and only about 1% of that found in wild-type pseudodiploid CHO cells, yet this low amount was sufficient to confer a selectable DHFR-positive growth phenotype. Thus, relatively minor perturbations in splicing are detectable in this system. The selection method does have an important restriction; these mutations must result in at least some skipping of the extra exon 2A, as it is the splicing of exon 1 to the remaining exon 2 that recreates the functional *dhfr* mRNA which is the basis of the selection (Fig. 1A). It is possible that this constraint has limited the spectrum of mutations observed. However, in general in vivo mutations at splice sites of internal exons do result in exon skipping (see below); in particular, five mutations that knock out the splicing of exon 2 in the genomic *dhfr* gene in CHO cells (three at the donor site and two at the acceptor site) all display exon 2 skipping (11).

Saturation mutagenesis. Our initial goal in these experiments was to use in vivo mutagenesis to saturate all possible nucleotide positions affecting splicing. To maximize the bases and sequences that could be hit, we used five different mutagenic treatments (including no mutagen). In theory, both A · T and G · C pairs should have been targeted. However, each of the four mutagens we used displayed surprisingly high site specificity.

Strand specificity is one factor that may contribute to the susceptibility of positions to mutagenesis. In several studies of mutation in cultured mammalian cells, it has been observed that the nontranscribed strand is much more prone to mutation by agents that produce bulky adducts or pyrimidine dimers (8, 16, 77, 81). Our data are also consistent with this finding. If we assume that BcPHDE, EMS, and ENU target purines and that UV targets pyrimidine dimers, then we can conclude that the great majority (84%) of the induced mutations occurred on the nontranscribed strand (the upper strand in Fig. 4). These results are not due to selection bias, since there are many potential sites on the transcribed (bottom) strand where mutation at that position does affect splicing (e.g., pyrimidine dimers at positions −2 to +1 of the acceptor site and at positions −3 to +1 of the donor site and at the branch point, and purines at positions −5 and −6 of the acceptor site).

However, the site specificity of these mutagens is also apparent within the nontranscribed strand. For example, seven of nine independent treatments with EMS yielded the same base change at the same position (Fig. 4). EMS is

expected to alkylate G residues (67) and is known to induce predominantly G · C to A · T transitions in mammalian cells (6, 42). Indeed, the two guanine bases at donor positions −1 and +1 on the nontranscribed strand (top strand in Fig. 4) were the probable targets. But it is not clear why the two guanines at positions −1 and +1 of the acceptor site were spared; these sites are necessary for efficient splicing and are vulnerable to chemical attack, as they were hit by other mutagens (Fig. 4). Similarly, BcPHDE is an aromatic hydrocarbon that can form adducts with A or G (7, 22). The action of this mutagen has been studied at the *dhfr* locus in CHO cells; both A and G were targets, with a great preference for the 5′ base of di- or tripurine sequences (13). Moreover, this mutagen displayed a striking propensity to hit splice acceptor sites. This specificity was also seen here; 12 of 20 BcPHDE-induced mutations were at the splice acceptor, where the affected purine was at the 5′ end of an oligopurine stretch. However, at the other major BcPHDE site, the branch point, the predominant change was at the 3′ position of a dipurine. In contrast, the sensitive and vulnerable tetrapurine sequence at positions −3 to +1 of the donor site was spared by BcPHDE. Similar preferential sites of action were seen with UV and ENU and for spontaneous mutation. These results suggest that DNA polymorphism may play a determining role in susceptibility to specific mutagens.

Whatever the cause, the fact that all these treatments generate hot and cold spots for mutation may have thwarted our efforts to hit all positions that are relevant for splicing. Thus, although this sensitive method produced 66 mutants at 16 sites, we cannot be sure that some positions necessary for splicing have been missed because of a low susceptibility to mutagenesis. We conclude that whereas in vivo saturation mutagenesis may be used to identify all genes in a pathway, it does not serve to identify all *cis*-acting positions at the nucleotide level. For this purpose the generation of mutations in synthetic oligonucleotides is probably required. This strategy did enable us to generate over 30 independent mutants affecting a specific splice donor site and a similar number affecting a specific splice acceptor site, many times more than what has been achieved by simply screening of negative mutants.

The 5′ donor splice site mutations. Donor site mutations included four different changes within the GU dinucleotide. AU, UU, and GA abolished correct splicing, but a GC dinucleotide (NS60) continued to splice correctly about 20% of the time. This result is in agreement with those of Aebi et al. (2, 3), who studied the effect of site-directed mutations on the splicing of rabbit β-globin pre-mRNA in a cell-free system or after transfection; they found the variant GC dinucleotide was the only one that allowed correct splicing. GC is by far the most common naturally occurring nonconsensus donor dinucleotide, with about 24 cases known (37). In all of these instances, the surrounding bases show greater agreement to the remainder of the consensus, as if to compensate for the weaker binding to U1 RNA that the GC sequence implies; in particular, the G at position +5 is 100% conserved. In the *dhfr* intron 2 donor site a U occupies the +5 position; perhaps correct splicing in this mutant would be higher than 20% were this position the consensus G.

Unlike the dinucleotide at positions +1 and +2, the remaining bases in the 9-base consensus are far from universally conserved. For the consensus donor sequence shown in Fig. 3, the occupancy by consensus bases is 71, 63, 82, 94, 74, 81, and 47% for positions −3, −2, −1, +3, +4, +5, and +6, respectively, in a rodent data base (64). However, this laxity in consensus agreement was not accompanied by leaky phenotypes for mutants altered at these positions. Mutations at −3, −2, −1, +3, and +4 almost always eliminated correct splicing (Table 1). Each of these mutations reduced the agreement of the splice site sequence with the consensus, supporting the idea that the overall match to the consensus is important (3, 66, 82). However, the interpretation of these results in terms of U1 RNA binding (87) is less clear (Table 1). For instance, at position −1, where the consensus G in the wild-type sequence can pair with C in U1 RNA, a G-to-A change eliminated correct splicing (NS8), whereas a G-to-U change allowed more than 30% correct splicing (NS92). And an A-to-G change at position +4 severely decreased correct splicing (NS90), even though the mutant G should be able to partially base pair with the pseudouridine at the corresponding position in U1. Thus, whereas it is clear that many positions are critical for activity at a particular donor site, the rationale for the exact sequence requirement for splicing at any given site remains elusive.

No mutants were found altered at the +5 position of the donor site. This result is not surprising, since the U at this position in the wild-type sequence already deviates from the consensus G. At position +6, a consensus U is present in the wild-type sequence, yet no mutations were found at this position either. This negative result is also understandable considering U1 RNA binding; in the absence of a G · C pair at position +5, the potential U · A pair at the end position (+6) cannot provide any stacking interaction and so its energy contribution is probably too low to be important. An examination of a data base of primate splice donors supports this argument; among splice donor sequences that lack a G at position +5, the occupancy of position +6 is nearly random (21% A, 19% C, 31% G, and 29% T, 180 sequences [36]).

Along with exon skipping, the activation of a cryptic splice site is a potential consequence of the destruction of a normal splice site (for a review, see reference 30). Little such activation was seen among the *dhfr* mutants. Within intron 2, there is a cryptic donor site located 31 bases downstream of the normal site; this site is used in a few mutants (Table 1). Its sequence, AGG/GUUUGU, has a modest fit to the consensus, 69.8 compared to 86.3 for the wild-type sequence (scoring system of Shapiro and Senapathy [65]), which may explain its low use. What is puzzling is why this cryptic site is activated by four donor site mutations (Table 1, NS7, 82, 60, and 90) but not by seven others (NS1, 79, 42, 92, 8, 5, and 91). This dichotomy extends to mutations at the same position; a G-to-A change at +1 activates the cryptic splice, whereas a G-to-U change does not. Computer-aided secondary structure analysis showed that the cryptic site can be part of a stem-loop structure. Although most of the splice donor mutations altered potential secondary structures in the vicinity of the normal splice site, they had no effect on the structure at the cryptic site. Perhaps some mutant sequences continue to bind splicing factors tightly and preclude their use nearby, while others bind factors poorly and allow them to scan locally for alternative sites.

One donor site mutation (NS1) fell outside the 9-base consensus sequence, 8 bases upstream of the donor site at the 3′ end of exon 2A. This position is outside the region that is known to base pair with U1 RNA. Exon mutations lying outside the splice consensus sequence have been shown to affect splicing in natural pre-mRNA molecules (32, 70, 71, 86) as well as in artificial constructs (24, 59, 73). These results have led to the idea that the exon contributes a favorable context for splicing (52), presumably through the formation of particular secondary or tertiary structures (19,

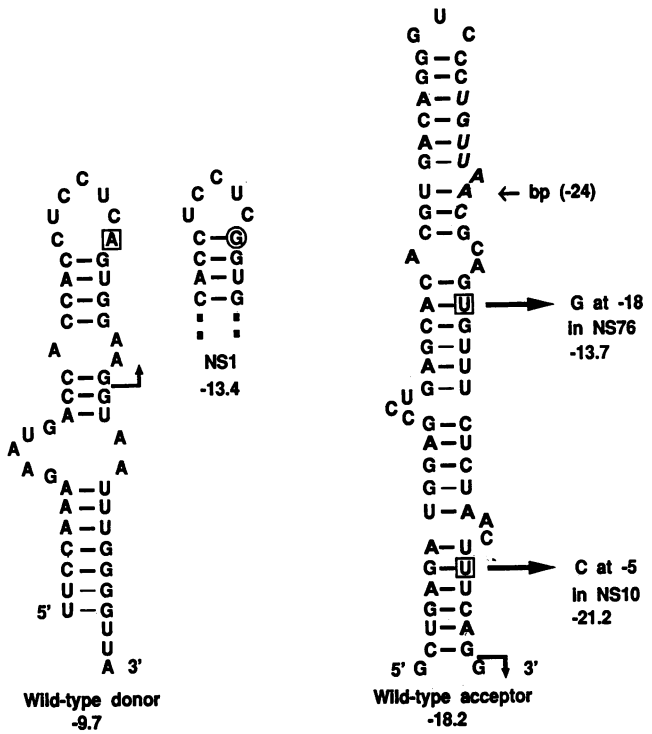

FIG. 7. Possible RNA secondary structural consequences of splicing mutations that do not disrupt consensus splice sequences. Structures were generated by using the MFOLD program of M. Zuker and J. Jaeger (version 2.0). Wild-type bases at mutated positions are boxed. G-U bonds are depicted with a thinner line. The bent arrows demarcate the exons, pointing in their direction. The numbers show the calculated free energies of binding. Left: sequence surrounding the donor site of exon 2A. The top portion of the homologous structure in mutant NS1 is also shown. Right: sequence flanking the acceptor site of exon 2A. The putative branch site sequence, matching the consensus, is italicized; bp, branch point. Numbers in parentheses are relative to the acceptor splice site.

21, 43). In this regard, we might have expected more than this single exon mutation to have appeared in our experiments, given the sensitivity of our selection method (capable of detecting 15% exon skipping). The fact that only one such mutation emerged argues that, if exon secondary structures do play a role, these structures are not easily disrupted by single-base changes.

An RNA secondary structure computed (89) for the 48-base region surrounding the exon 2 donor site yielded the stem-loop structure shown in Fig. 7 (left), with the splice site located in one strand of the stem. The stability of this stem was enhanced more than 3 kcal/mol in mutant NS1 pre-mRNA (Fig. 7); perhaps this increased stability interferes with splicing by hindering the unwinding of the double-stranded structure.

The 3′ splice acceptor site. At the splice acceptor site, mutations were found at 3 of the 4 consensus positions that surround the splice acceptor site (CAG/G). A change from G to U at the first base in the exon produced a leaky phenotype, with half of the transcripts continuing to splice at this site. This weak phenotype correlates with the fact that this +1 position is relatively poorly conserved (52% [64]), with U being the least common occupant (9%). This same change at the acceptor site preceding exon 5 in the genomic *dhfr* gene leads to essentially 100% exon skipping (47), underscoring the idiosyncracy of individual splice site requirements. Mutations at the near universally conserved AG completely obliterated correct splicing. In some mutants the amount of mRNA that accumulated was substantially reduced (e.g., NS3 and NS75 [Table 1]), perhaps because many transcripts may have been blocked at the level of unstable splicing intermediates.

The wild-type polypyrimidine tract at this acceptor site contains 9 of 11 pyrimidines. Substitutions of a purine for U at position −5 or −6 reduced this ratio to 8 of 11 and reduced correct splicing by 50 to 90%. The consequences are actually more extreme than the lower figure implies; if the lower overall amount of *dhfr* mRNA is taken into account, the correctly spliced mRNA represents only 5 to 15% of the wild-type level (Table 1). Such substantial effects for single-base substitutions in the polypyrimidine tract are rare for constitutively spliced exons (5).

More surprising was the deleterious effect of substituting one pyrimidine for another at position −5. Here a change of U to C decreased correct splicing by half (mutant NS10), suggesting that U may play a more specific role than is apparent from its slight plurality at this position in the rodent consensus (49 versus 38% [64]). A preference for U in the polypyrimidine stretch has also been noted in yeasts (56) and has been suggested by the ineffectiveness of a polypyrimidine C in improving the splicing of an especially short exon (23). As in the case of the exon mutation NS1 described above, the effect of this pyrimidine-for-pyrimidine switch may be interpreted in terms of RNA secondary structure. On the right side of Fig. 7 is shown the computer-generated stem-loop encompassing the region flanking the acceptor site. The 3 kcal/mol stabilization of this structure in NS10 may hinder access to the acceptor site.

A pyrimidine-to-purine change at position −18 also reduced the proportion of correct splicing by half (mutant NS76). This position lies about halfway between the branch point consensus starting at −23 and the usual extent of the polypyrimidine tract (to −14 or −15), and the mutation may be affecting either element. As seen in Fig. 7 (right), the base substitution in NS76 destabilizes the secondary structure predicted for the wild-type sequence and, by the arguments given above, should enhance accessibility of this region to splicing factors. It is possible that this sequence change is acting more directly (see below). Alternatively, this change may be creating a potential alternative splice acceptor ($Y_{8/11}$NCAG/G) that distracts the splicing machinery.

The branch point. The branch point used for lariat formation in intron 1 of *dhfr* pre-mRNA has not been experimentally determined. However, the prevalence of mutations at position −24 and the fact that this position coincides with position 6 (the branch position) of a 7-base consensus branch point sequence (YNYURAY [53]) makes a compelling argument that this A is indeed the branch point. Mutation of the branch nucleotide often has little effect on the splicing of mammalian exons, with branch formation occurring at nearby alternative sites (34, 62). In contrast, the mutation of the *dhfr* intron 1 branch point seen here produced a very substantial exon skipping phenotype, at least for an A-to-G substitution (NS14 [Table 1]). Substitution of U had a lesser effect, although skipping still predominated. The relative effects of the two types of substitution agree with the observation that lariat formation can occur more readily at a branch point U than a G (34). One mutation was found at a consensus position other than the branch nucleotide itself; this A-to-U change at position −25 (NS13) had only a small

effect on splicing, despite the weaker binding to U2 RNA that this change implies (85, 88). A similar negative result was found for this same change in in vitro studies of a β-globin branch point (60). The marginal phenotype of mutant NS13 did help define the sensitivity of the selection.

The branch point sequence surrounding position −24 is one of two 7-mers in intron 1 that conform to the consensus sequence. However, the other 7-mer would place the branch point at −8, which may be too close to the acceptor site. There are several other sites that contain 6 consensus bases (e.g., at positions −40 and −60). Since some inclusion of exon 2A is seen in all of the branch point mutants, it is possible that branch points are being formed at these or other alternative sites in addition to or instead of the normal site.

Mutant NS76 carries a U-to-G change at a position 6 bases downstream of the end of the putative branch point, a position well beyond the branch point consensus. Reed (58) has shown that the pyrimidine content near the branch point is important, as the replacement of multiple pyrimidines by purines in this region reduced splicing; the NS76 mutation supports this conclusion. Moreover, none of the 31 authentic mammalian branch site sequences compiled by Nelson and Green (53) has a G at position +6 relative to the branch nucleotide. Thus, it may be that the G at this position in mutant NS76 is directly interfering with binding of a branch site factor. In fact, Penotti (57) has suggested an extended 21-bp branch point consensus sequence, stretching from −5 to +15 relative to the branch nucleotide, and often overlapping the polypyrimidine region. This extended consensus sequence does not display an extended complementarity to U2 RNA.

Mutations in all three splice consensus sequences lead to exon skipping. We have previously shown that mutation at the splice sites bordering exon 5 (47) and exon 2 (12) of the hamster *dhfr* gene cause the skipping of the exon, rather than the activation of cryptic splice sites. Many examples demonstrating the generality of this phenotype for diverse genes have accumulated over the past few years, mainly from the analysis of human genetic disease (references too numerous to cite, but for some donor site examples see references 14, 33, and 40). Moreover, we have found that exon skipping is the predominant phenotype for induced mutations at all four internal exons of the *dhfr* gene in CHO cells (11), and similar results have been found for several exons of the hypoxanthine phosphoribosyltransferase gene (70, 86).

The results reported here also extend the generality of the exon skipping consequence. It is true that in our experiments we have selected for this phenotype. Moreover, it is possible that some unspliced or aberrantly spliced RNA molecules do result from these mutations, but we have failed to detect them because they are rapidly degraded or because they fail to amplify efficiently in the PCR. However, molecules in which exon 2A is skipped do accumulate substantially in most mutants, to more than 50% of the parental mRNA level in the majority of mutants analyzed. We conclude that mutation at most, if not all, of the positions known to affect splicing can lead to the skipping of the affected exon.

Most previous mutations shown to effect exon skipping of constitutive exons have been at splice donor or acceptor sites. That mutations at the branch site can also induce exon skipping is consistent with results for alternatively spliced transcripts, where a weak or weakened branch site has been shown to be an important element in the choice between two splice acceptor sites (1, 26, 54, 68).

Exon skipping is a phenotype predicted by the exon definition model of splice site selection. Using site-directed mutagenesis followed by cell-free and transfected-cell analysis, Berget and her colleagues showed that mutations at a donor site could reduce splicing at the upstream acceptor site (61, 72). They proposed a model in which the two ends of the exon must be recognized by the splicing machinery as a prerequisite for subsequent splicing events. The multiplicity of mutations found here that result in the skipping of exon 2A lends strong support to this exon definition model.

Splicing depends on a small number of specific sites. Five different treatments of NB6 cells produced 66 splicing mutations; these 66 mutations were concentrated at only eight 5′ and eight 3′ positions surrounding exon 2. This limited spectrum can be viewed in two ways. The first interpretation acknowledges that the splice consensus sequence does not contain sufficient information to demarcate a splice site and assumes that there are many other specific bases that contribute to the definition of a particular site. Why then weren't these positions unmasked by mutations in this system, a system that could detect a mere 15% splicing irregularity? The limited mutational spectrum could be ascribed to the high sequence specificity of the mutagens and the imposed requirement for exon skipping. We argue against this interpretation, because we see no reason to think that the mutagenic treatments should have spared important bases outside the consensus sequence and no reason to think that the phenotypic consequence of exon skipping should be confined to consensus sequence positions.

We favor a second interpretation, one that questions the requirement for specific additional sequences to define a particular splice site. In this view, splice site recognition would depend on a context (52) of RNA secondary and tertiary structure. The interactions producing this context would be sufficiently robust to be immutable by most single-base changes.

Three mutations that did not disrupt a conventional branch point, acceptor, or donor consensus sequence were indeed isolated: a base substitution at −18 of the acceptor site, a U-to-C pyrimidine substitution at −5 of the acceptor site, and a base substitution at −8 of the donor site. These three leaky mutations may represent clues for further probing of the structural determinants of splice site definition.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant GM-22629 from the National Institute of General Medical Sciences. I-T.C. was supported by National Cancer Institute Research Fellowship CA08844.

REFERENCES

1. **Adema, G. J., K. L. van Hulst, and P. D. Baas.** 1990. Uridine branch acceptor is a cis-acting element involved in regulation of the alternative processing of calcitonin/CGRP-I pre-mRNA. Nucleic Acids Res. **18**:5365–5373.
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. **Aebi, M., H. Hornig, and C. Weissmann.** 1987. 5′ cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5′ splice region, not by the conserved 5′GU. Cell **50**:237–246.
4. **Aebi, M., and C. Weissmann.** 1987. Precision and orderliness in splicing. Trends Genet. **3**:102–107.
5. **Archibald, A. L., N. A. Thompson, and S. Kvist.** 1986. A single nucleotide difference at the 3′ end of an intron causes differential splicing of two histocompatibility genes. EMBO J. **5**:957–965.
6. **Ashman, C. R., and R. L. Davidson.** 1987. DNA base sequence

changes induced by ethyl methanesulfonate in a chromosomally integrated shuttle vector gene in mouse cells. Somatic Cell Mol. Genet. **13**:563–568.

7. **Bigger, C., J. Strandberg, H. Yagi, D. Jerina, and A. Dipple.** 1989. Mutagenic specificity of a potent carcinogen, benzo[c]phenanthrene (4R,3S)α-dihydrodiol (2S,1R)α-epoxide, which reacts with adenine and guanine in DNA. Proc. Natl. Acad. Sci. USA **86**:2291–2295.
8. **Carothers, A. M., J. Mucha, and D. Grunberger.** 1991. DNA strand-specific mutations induced by (±)-3α,4β-dihydroxy-1α,2α-epoxy-1,2,3,4-tetrahydrobenzo[c]phenanthrene in the dihydrofolate reductase gene. Proc. Natl. Acad. Sci. USA **88**: 5749–5753.
9. **Carothers, A. M., G. Urlaub, N. Ellis, and L. A. Chasin.** 1983. Structure of the dihydrofolate reductase gene in Chinese hamster ovary cells. Nucleic Acids Res. **11**:1997–2012.
10. **Carothers, A. M., G. Urlaub, D. Grunberger, and L. A. Chasin.** 1988. Mapping and characterization of mutations induced by benzo[a]pyrene diol epoxide at the dihydrofolate reductase locus in CHO cells. Somatic Cell Mol. Genet. **14**:169–183.
11. **Carothers, A. M., G. Urlaub, D. Grunberger, and L. A. Chasin.** 1992. Unpublished data.
12. **Carothers, A. M., G. Urlaub, D. Mucha, D. Grunberger, and L. A. Chasin.** 1989. Point mutation analysis in a mammalian gene: rapid preparation of total RNA, PCR amplification of cDNA, and Taq sequencing by a novel method. BioTechniques **7**:494–499.
13. **Carothers, A. M., G. Urlaub, J. Mucha, R. Harvey, L. A. Chasin, and D. Grunberger.** 1990. Splicing mutations in the CHO DHFR gene preferentially induced by (±)-3α,4β-dihydroxy-1α,2α-epoxy-1,2,3,4,-tetrahydrobenzo[c]phenanthrene. Proc. Natl. Acad. Sci. USA **87**:5464–5468.
14. **Carstens, R. P., W. A. Fenton, and L. R. Rosenberg.** 1991. Identification of RNA splicing errors resulting in human ornithine transcarbamylase deficiency. Am. J. Hum. Genet. **48**: 1105–1114.
15. **Chen, I-T., and L. A. Chasin.** 1992. Update on point mutation analysis in a mammalian gene, p. 60. *In* J. Ellingboe and U. B. Gyllensten (ed.), The PCR technique: DNA sequencing. Eaton Publishing Co., Natick, Mass.
16. **Chen, R. W., V. M. Maher, and J. J. McCormick.** 1990. Effect of excision repair by diploid human fibroblasts on the kinds and locations of mutations induced by BPDE in the coding region of the HPRT gene. Proc. Natl. Acad. Sci. USA **87**:8680–8684.
17. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
18. **Ciudad, C. J., G. Urlaub, and L. Chasin.** 1988. Deletion analysis of the Chinese hamster dihydrofolate reductase gene promoter. J. Biol. Chem. **263**:16274–16282.
19. **Clouet-d'Orval, B., Y. d'Aubenton-Carafa, P. Sirand-Pugnet, M. Gallego, E. Brody, and J. Marie.** 1991. RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. Science **252**:1823–1828.
20. **Crouse, G. F., R. N. McEwan, and M. L. Pearson.** 1983. Expression and amplification of engineered mouse dihydrofolate reductase minigenes. Mol. Cell. Biol. **3**:257–266.
21. **Deshler, L. O., and J. J. Rossi.** 1991. Unexpected point mutations activate cryptic 3′ splice sites by perturbing a natural secondary structure within a yeast intron. Genes Dev. **5**:1252–1263.
22. **Dipple, A., M. A. Pigott, S. K. Agarwal, H. Yagi, J. M. Sayer, and D. M. Jerina.** 1987. Optically active benzo[c]phenanthrene diol epoxides bind extensively to adenine in DNA. Nature (London) **327**:535–536.
23. **Dominski, Z., and R. Kole.** 1991. Selection of splice sites in pre-mRNAs with short internal exons. Mol. Cell. Biol. **11**:6075–6083.
24. **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.
25. **Freier, S., R. Kierzek, J. A. Jaeger, N. Sugimoto, M. Caruthers, T. Neilson, and D. H. Turner.** 1986. Improved free-energy parameters for predictions of RNA duplex stability. Proc. Natl. Acad. Sci. USA **83**:9373–9377.
26. **Fu, X.-D., R. A. Katz, A. M. Skalka, and T. Maniatis.** 1991. The role of branchpoint and 3′-exon sequences in the control of balanced splicing of avian retrovirus RNA. Genes Dev. **5**:211–220.
27. **Fu, X.-D., and T. Maniatis.** 1990. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. Nature (London) **343**:437–441.
28. **Fu, X.-Y., and J. L. Manley.** 1987. Factors influencing alternative splice site utilization in vitro. Mol. Cell. Biol. **7**:738–748.
29. **Fuscoe, J. C., R. G. Fenwick, D. H. Ledbetter, and C. T. Caskey.** 1983. Deletion and amplification of the HGPRT locus in Chinese hamster cells. Mol. Cell. Biol. **3**:1086–1096.
30. **Green, M. R.** 1986. Pre-mRNA splicing. Annu. Rev. Genet. **20**:671–708.
31. **Green, M. R.** 1991. Biochemical mechanisms of constitutive and regulated pre-mRNA splicing. Annu. Rev. Cell Biol. **7**:559–599.
32. **Hampson, R., L. Follette, and F. Rottman.** 1989. Alternative processing of bovine growth hormone mRNA is influenced by downstream exon sequences. Mol. Cell. Biol. **9**:1604–1610.
33. **Hidaka, Y., T. D. Palella, T. E. O'Toole, S. A. Tarlé, and W. N. Kelley.** 1987. Human adenine phosphoribosyltransferase: identification of allelic mutations at the nucleotide level as a cause of complete deficiency of the enzyme. J. Clin. Invest. **80**:1409–1415.
34. **Hornig, H., M. Aebi, and C. Weissmann.** 1986. Effect of mutations at the lariat branch acceptor site on β-globin pre-mRNA splicing in vitro. Nature (London) **324**:589–591.
35. **Huang, S., and D. L. Spector.** 1991. Nascent pre-mRNA transcripts are associated with nuclear regions enriched in splicing factors. Genes Dev. **5**:2288–2302.
36. **Ip, J., G. Landau, J. Schmidt, and L. A. Chasin.** 1992. Unpublished data.
37. **Jackson, I. L.** 1991. A reappraisal of non-consensus mRNA splice sites. Nucleic Acids Res. **19**:3795–3798.
38. **Kedes, D. H., and J. A. Steitz.** 1987. Accurate 5′ splice-site selection in mouse kappa-immunoglobin light chain premessenger RNAs is not cell-type specific. Proc. Natl. Acad. Sci. USA **84**:7928–7932.
39. **Kedes, D. H., and J. A. Steitz.** 1988. Correct in vivo splicing of the mouse immunoglobulin κ light-chain pre-mRNA is dependent on the 5′ splice-site position even in the absence of transcription. Genes Dev. **2**:1448–59.
40. **Kuivaniemi, H., S. Kontusaari, G. Tromp, M. Zhao, C. Sabol, and D. J. Prockop.** 1990. Identical G^{+1} to A mutations in three different introns of the type III procollagen gene (COL3A1) produce different patterns of RNA splicing in three variants of Ehlers-Danlos syndrome IV. J. Biol. Chem. **265**:12067–12074.
41. **Lawrence, J. B., R. H. Singer, and L. M. Marselle.** 1989. Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. Cell **57**:493–502.
42. **Lebkowski, J., J. H. Miller, and M. P. Calos.** 1986. Determination of DNA sequence changes induced by ethyl methanesulfonate in human cells using a shuttle vector system. Mol. Cell. Biol. **6**:1838–1842.
43. **Libri, D., A. Piseri, and M. Y. Fiszman.** 1991. Tissue-specific splicing in vivo of the β-tropomyosin gene: dependence on an RNA secondary structure. Science **252**:1842–1845.
44. **Lossi, A.-M., and J.-L. Berge-Lefranc.** 1989. The mRNA transcripts from a mutant β-globin gene derived from splicing at preferential cryptic sites. FEBS Lett. **256**:163–166.
45. **Lowery, D., and B. Van Ness.** 1987. In vitro splicing of kappa immunoglobulin precursor mRNA. Mol. Cell. Biol. **7**:1346–1351.
46. **Lowy, I., A. Pellicer, J. F. Jackson, G.-K. Sim, S. Silverstein, and R. Axel.** 1980. Isolation of transforming DNA: cloning the hamster aprt gene. Cell **22**:817–823.
47. **Mitchell, P., G. Urlaub, and L. A. Chasin.** 1986. Spontaneous splicing mutations at the dihydrofolate reductase locus in CHO cells. Mol. Cell. Biol. **6**:1926–1935.

48. **Mitchell, P. J., A. M. Carothers, J. H. Han, J. D. Harding, E. Kas, L. Venolia, and L. A. Chasin.** 1986. Multiple transcription start sites, DNase I-hypersensitive sites, and an opposite strand exon in the 5′ region of the CHO dhfr gene. Mol. Cell. Biol. **6**:425–440.
49. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
50. **Mulligan, R. C., and P. Berg.** 1980. Expression of a bacterial gene in mammalian cells. Science **209**:1422–1427.
51. **Nalbantoglu, J., D. Goncalves, and M. Meuth.** 1985. Structure of mutant alleles at the APRT locus of CHO cells. J. Mol. Biol. **167**:575–594.
52. **Nelson, K. K., and M. R. Green.** 1988. Splice site selection and ribonucleoprotein complex assembly during in vitro pre-mRNA splicing. Genes Dev. **2**:319–329.
53. **Nelson, K. K., and M. R. Green.** 1989. Mammalian U2 snRNA has a sequence-specific RNA-binding activity. Genes Dev. **3**:1562–1571.
54. **Noble, J. C., C. Prives, and J. Manley.** 1988. Alternative splicing of SV40 early pre-mRNA is determined by branch site selection. Genes Dev. **2**:1460–75.
55. **Orkin, S. H., and H. H. Kazazian, Jr.** 1984. The mutation and polymorphism of the human β-globin gene and its surrounding DNA. Annu. Rev. Genet. **18**:131–71.
56. **Patterson, B., and C. Guthrie.** 1991. A U-rich tract enhances usage of an alternative 3′ splice site in yeast. Cell **64**:181–187.
57. **Penotti, F. E.** 1991. Human pre-mRNA splicing signals. J. Theor. Biol. **150**:385–420.
58. **Reed, R.** 1989. The organization of 3′ splice-site sequences in mammalian introns. Genes Dev. **3**:2113–2123.
59. **Reed, R., and T. Maniatis.** 1986. A role for exon sequences and splice-site proximity in splice-site selection. Cell **46**:681–690.
60. **Reed, R., and T. Maniatis.** 1988. The role of the mammalian branchpoint sequence in pre-mRNA splicing. Genes Dev. **2**:1268–1276.
61. **Roberson, B., G. Cote, and S. Berget.** 1990. Exon definition may facilitate splice site selection in RNAs with multiple exons. Mol. Cell. Biol. **10**:84–94.
62. **Ruskin, B., J. Greene, and M. Green.** 1985. Cryptic branch point activation allows accurate in vitro splicing of human β-globin intron mutants. Cell **41**:833–44.
63. **Seetharam, S., and I. B. Dicker.** 1991. A rapid and complete 4-step protocol for obtaining nucleotide sequence from E. coli genomic DNA from overnight cultures. BioTechniques **11**:32–34.
64. **Senapathy, P., M. P. Shapiro, and N. Harris.** 1990. Splice junctions, branch point sites, and exons: sequence statistics, identification, and applications to genome project. Methods Enzymol. **183**:252–278.
65. **Shapiro, M. P., and P. Senapathy.** 1987. RNA splice junction of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res. **15**:7155–7174.
66. **Siliciano, P., and C. Guthrie.** 1988. 5′ splice site selection in yeast: genetic alterations in base-pairing with U1 reveal additional requirements. Genes Dev. **2**:1258–1267.
67. **Singer, B., and D. Grunberger.** 1983. Molecular biology of mutagens and carcinogens, p. 56–59. Plenum Press, New York.
68. **Smith, C. W. J., and B. Nadal-Ginard.** 1989. Mutually exclusive splicing of α-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. Cell **56**:749–758.
69. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
70. **Steingrimsdottir, H., G. Rowley, G. Dorado, J. Cole, and A. R. Lehmann.** 1992. Mutations which alter splicing in the human hypoxanthine guanine phosphoribosyltransferase gene. Nucleic Acids Res. **20**:1201–1208.
71. **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.
72. **Talerico, M., and S. Berget.** 1990. Effect of 5′ splice site mutations on splicing of the preceding intron. Mol. Cell. Biol. **10**:6299–6305.
73. **Turnbull-Ross, A. D., A. J. Else, and I. C. Eperon.** 1988. The dependence of splicing efficiency on the length of 3′ exon. Nucleic Acids Res. **16**:395–411.
74. **Turner, D. H., N. Sugimoto, and S. Freier.** 1988. RNA structure prediction. Annu. Rev. Biophys. Chem. **17**:167–192.
75. **Urlaub, G., and L. A. Chasin.** 1980. Isolation of Chinese hamster cell mutants deficient in dihydrofolate reductase activity. Proc. Natl. Acad. Sci. USA **77**:4216–4220.
76. **Urlaub, G., E. Kas, A. M. Carothers, and L. A. Chasin.** 1983. Deletion of the diploid dihydrofolate reductase locus from cultured mammalian cells. Cell **33**:405–412.
77. **Urlaub, G., P. J. Mitchell, C. J. Ciudad, and L. A. Chasin.** 1989. Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. Mol. Cell. Biol. **9**:2868–2880.
78. **Urlaub, G., P. J. Mitchell, E. Kas, L. A. Chasin, V. L. Funanage, T. T. Myoda, and J. L. Hamlin.** 1986. The effect of gamma rays at the dihydrofolate reductase locus: deletions and inversions. Somatic Cell Mol. Genet. **12**:555–566.
79. **Venolia, L., G. Urlaub, and L. A. Chasin.** 1987. Polyadenylation of Chinese hamster dihydrofolate reductase genomic genes and minigenes after gene transfer. Somat. Cell Mol. Genet. **13**:491–501.
80. **Vrieling, H., M. J. Niericker, J. W. I. M. Simons, and A. A. van Zeeland.** 1988. Molecular analysis of mutations induced by N-ethyl-N-nitrosourea at the HPRT locus in mouse lymphoma cells. Mutat. Res. **198**:99–106.
81. **Vrieling, H., M. L. Van Rooouen, N. A. Groden, M. Z. Zdzienicka, J. W. I. M. Simons, P. H. Lohman, and A. A. van Zeeland.** 1989. DNA strand specificity for UV-induced mutations in mammalian cells. Mol. Cell. Biol. **9**:1277–1283.
82. **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′ terminal of U1 snRNA. Nucleic Acids Res. **16**:471–486.
83. **Wieringa, B., F. Meyer, J. Reiser, and C. Weissmann.** 1983. Unusual splice sites revealed by mutagenetic inactivation of an authentic splice site of the rabbit β-globin gene. Nature (London) **301**:38–43.
84. **Wigler, M., A. Pellicer, S. Silverstein, R. Axel, G. Urlaub, and L. A. Chasin.** 1979. Transformation of the APRT locus in mammalian cells. Proc. Natl. Acad. Sci. USA **76**:1373–1376.
85. **Wu, J., and J. L. Manley.** 1989. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. Genes Dev. **3**:1553–1561.
86. **Zhang, L.-H., H. Vrieling, A. A. van Zeeland, and D. Jenssen.** 1992. Spectrum of spontaneously occurring mutations in the hprt gene of V79 Chinese hamster cells. J. Mol. Biol. **223**:627–635.
87. **Zhuang, Y., and A. Weiner.** 1986. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. Cell **46**:827–835.
88. **Zhuang, Y., and A. Weiner.** 1989. A compensatory base change in human U2 snRNA can suppress a branch site mutation. Genes Dev. **3**:1545–1552.
89. **Zuker, M.** 1989. On finding all sub-optimal foldings of an RNA molecule. Science **244**:48–52.