A Novel Bipartite Intronic Splicing Enhancer Promotes the Inclusion of a Mini-exon in the AMP Deaminase 1 Gene*

Received for publication, December 22, 2000, and in revised form, April 26, 2001
Published, JBC Papers in Press, April 30, 2001, DOI 10.1074/jbc.M011637200

Thomas Genetta‡, Hiroko Morisaki§, Takayuki Morisaki§, and Edward W. Holmes‡‡

The Journal of Biological Chemistry Vol. 276, No. 27, Issue of July 6, pp. 25589–25597, 2001

Alternative splicing of the 12-base exon 2 of the adenosine monophosphate deaminase (AMPD) gene is subject to regulation by both cis- and trans-regulatory signals. The extent of exon 2 inclusion is stage- and cell type-specific and is subject to the physiological state of the cell. In adult skeletal muscle, a cell type that regulates the activity of this allosteric enzyme at several levels, the exon 2-plus form of AMPD, predominates. We have performed a systematic analysis of the cis-acting regulatory sequences that reside in the intron immediately downstream of this mini-exon. A complex element comprising sequences that enhance exon 2 inclusion and sequences that counteract this effect resides in the middle of this intron. We demonstrate that the enhancing component is bipartite, with more than a kilobase of sequence separating the two functional sites. The presence of even minimal levels the mini-exon in the fully processed AMPD mRNA requires both of these sites, neither of which appears in any other published splicing enhancer. An RNA binding activity derived from a muscle cell line requires both of the enhancing sites. Mutations in either of the sites that eliminate exon 2 inclusion abrogate this binding activity.

The vast majority of metazoan genes contain short, information-encoding exons interspersed by relatively long stretches of noncoding introns. The processing of this information to yield a translatable message, including the splicing of the pre-mRNA, is subject to regulation at virtually every definable step. One such step, including or excluding a particular exon—alternative splicing—provides a way to alter the functional activity of a protein in the absence of gene duplication (1, 2). Discreet functional domains may be added or subtracted depending on signals elaborated by the particular needs of the cell. The study of alternative splicing further provides an experimental tool for understanding how the participating ribonucleoprotein complexes target intron-exon boundaries (3).

Vertebrate exons are, on average, less than 300 nucleotides, whereas many introns are thousands of nucleotides in length (4). Exons below an average size of 50 nucleotides have been shown to be inherently difficult to recognize by the splicing machinery (4, 5). One explanation for this "masking" of small exons, originally set forth by Berget (5), is that the initiation of the splicing reaction is exon-centered. There is much evidence to support this model, termed exon definition, in which ribonucleoprotein initiation complexes recognize intron-exon boundaries and bridge across the exon. As exon size decreases below 50 nucleotides, these complexes are prevented from forming a productive interaction with the 3'- and 5'-splice recognition sites on the pre-mRNA or with each other, through steric hindrance. Sequences that function either to facilitate (enhancers) or to inhibit (repressors) the recognition of alternatively spliced exons have been found in both introns (6-9) and exons (9-11).

Adenosine monophosphate deaminase (AMPD) catalyzes a key step in purine nucleotide metabolism in virtually all eukaryotes. The purine nucleotide cycle serves as the sole source (in the form of fumarate) of citric acid cycle intermediates in contracting muscle tissue. A deficiency in AMPD is the most prevalent genetic disease in humans, the number of people heterozygous approaching 10% of Caucasian and individuals of African descent (12, 13). A small percentage of homozygous deficient individuals, nearly 2% of the affected populations, display symptoms of chronic fatigue and lost productivity as well as a greater predisposition to stress-related ailments, including heart disease and stroke (12, 14). Interestingly, a mutation in at least one AMPD allele appears to confer a protective effect on individuals at risk for one of the most prevalent diseases of industrialized nations, congestive heart failure. We have found that people harboring at least one AMPD mutant allele have a significantly prolonged probability of survival after the onset of symptoms leading to this extremely serious medical condition (15).

In muscle cells, AMPD is fully activated only when bound to myosin heavy chain through its carboxyl terminus. This enzyme is also influenced allosterically by cellular levels of purine nucleotides, through a binding site near its amino terminus (12, 24-28). We have shown that the four-amino acid peptide encoded by exon 2 influences both of these properties in a fiber type-dependent fashion (19). In adult, fast twitch, glycolytic myofibers, the sensitivity of AMPD to cellular ATP/GTP levels is altered significantly by the presence of the exon 2 domain (14, 16, 19). In the resting state, ATP levels in the myocyte are relatively high, and a purine nucleotide binding site near the carboxyl terminus of AMPD is occupied. The exon 2-minus isoform of AMPD, which predominates in slow twitch, oxidative myofibers, can bind to myosin heavy chain and become activated in the presence of relatively high concentrations of ATP (12, 19). Thus, the essential role of AMPD in generating both

* This work was supported in part by National Institutes of Health Grant DK12314 (to E. W. H.). 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.

‡ To whom correspondence should be addressed: Office of the Dean, School of Medicine, University of California at San Diego, Rm. 1313 Basic Science Bldg., 9500 Gilman Dr., La Jolla, CA 92039-0602.

§ Recipient of a fellowship from the American Heart Association.

¶ Department of Biological Sciences, National Cardiovascular Center Research Institute, Suita, Osaka 565-0871, Japan, and School of Medicine, University of California at San Diego, La Jolla, California 92039-0602.

‡‡ Joseph P. Stokes Research Institute, Children’s Hospital of Philadelphia, Philadelphia, Pennsylvania 19104.

The abbreviations used are: AMPD, adenosine monophosphate deaminase; kb, kilobase(s); ExRE, exon retention element; RT-PCR, reverse transcription-polymerase chain reaction; bp, base pair(s).
We present in this report the initial characterization of sequences residing in the ExRE of intron 2 of the AMPD gene which regulate the extent of exon 2 inclusion in the final splicing product. We have narrowed the ExRE to two short, novel, discreet sequences in the middle of intron 2, separated by ~1,150 bases. We found that this bipartite splicing enhancer functions in an orientation- and sequence-dependent manner. In addition, there is an absolute requirement for both of these enhancing elements in exon 2 inclusion. We go on to demonstrate the presence of a myocyte-specific factor that is detected only when both of the enhancing sites are present together in a binding reaction.

**Experimental Procedures**

**AMPD Mini-gene Expression Constructs**

Manipulations, deletions, and mutations of intron 2 of the human AMPD gene were made according to standard recombinant DNA protocols and procedures (21). All constructs were sequenced by the University of Pennsylvania Medical Center DNA Sequencing Facility. The resulting intron 2 enhancer mutations were cloned into the \( \beta \)-actin-based expression vector (22) depicted in Fig. 1A.

**Cell Lines and Transfections**

Balb/c 3T3 fibroblasts and murine Soleus 8 myoblasts were maintained in Dulbecco's modified Eagle's medium plus 10% fetal calf serum supplemented with glutamine in an atmosphere containing 5% CO\(_2\). Cells were transfected using a standard calcium phosphate procedure (21). After 48, transiently transfected cells were harvested, and total RNA was isolated using a guanidinium-based procedure (21). Two sets of primers, which bound to exon 3) in the

**Qualitative RT-PCR Analysis**

Preparation of total RNA and reverse transcription reaction conditions for the polymerase chain reaction and subsequent gel analysis of PCR products were carried out as described (20). Two sets of primers, both specific for the human AMPD homolog, were used in these studies. PCR products of 69 bp (exon 2-plus) and 57 bp (exon 2-minus) were generated as described (20). RT-PCR analysis yielding products of 216 bp (exon 2-plus) or 204 (exon 2-minus) was generated using the primer H6 (5'-GTCTGGATCTCATCCACATC-3', which bound to exon 3) in the for use as a citric acid cycle intermediate) and IMP (from the deamination of AMP) can be retained in both fiber types, at least in part, by regulating the alternative splicing of exon 2.

The AMPD gene is comprised of 16 exons and is regulated both transcriptionally (13) and post-transcriptionally (16–18) in a development- and tissue-specific manner. The primary transcript is alternatively spliced with the exon 2-minus form predominating in all cells prenatally and in myoblasts postnatally. This 12-base mini-exon is largely, but not exclusively, retained in adult myotubes (17). Recent data suggest a role for the exon 2-encoded peptide in altering the allosteric responsiveness of AMPD to intracellular ATP levels, clearly important for the homeostasis of muscle (19). A recently discovered genetic lesion, a C-T transition that results in the introduction of a nonsense mutation at the end of exon 2, is partly responsible for human AMPD deficiency. As a consequence of alternative splicing, however, this mutated exon is excluded in 0.6–2% of the enzyme in adult muscle, resulting in a partial rescue of the deficiency. Most individuals harboring this mutation are therefore asymptomatic (14). For all of these reasons, we are very interested in the molecular mechanisms by which the alternative splicing of AMPD is regulated.

We have demonstrated recently that alternative splicing of exon 2 is driven largely by two competing factors (20). First, the short distance between the suboptimal 5'-donor and 3'-acceptor splice recognition sites of this 12-base exon makes its recognition by the splicing apparatus inherently difficult. Second, sequences located roughly in the middle of the 5.2-kb downstream intron, the exon retention element (ExRE), are required for inclusion of the this exon in the final splicing product. We have shown that the strength of these two opposing influences on exon 2 inclusion is influenced greatly by cell type. In non-muscle cells, such as fibroblasts, exon 2 is included in slightly less than half of the final splice products. In myoblasts, the balance is shifted dramatically toward inclusion, about 90%, and differentiated myotubes push this even further, to a 97% inclusion rate (20).

We present in this report the initial characterization of sequences residing in the ExRE of intron 2 of the AMPD gene which regulate the extent of exon 2 inclusion in the final splicing product. We have narrowed the ExRE to two short, novel, discreet sequences in the middle of intron 2, separated by ~1,150 bases. We found that this bipartite splicing enhancer functions in an orientation- and sequence-dependent manner. In addition, there is an absolute requirement for both of these enhancing elements in exon 2 inclusion. We go on to demonstrate the presence of a myocyte-specific factor that is detected only when both of the enhancing sites are present together in a binding reaction.

**Experimental Procedures**

**AMPD Mini-gene Expression Constructs**

Manipulations, deletions, and mutations of intron 2 of the human AMPD gene were made according to standard recombinant DNA protocols and procedures (21). All constructs were sequenced by the University of Pennsylvania Medical Center DNA Sequencing Facility. The resulting intron 2 enhancer mutations were cloned into the \( \beta \)-actin-based expression vector (22) depicted in Fig. 1A.

**Cell Lines and Transfections**

Balb/c 3T3 fibroblasts and murine Soleus 8 myoblasts were maintained in Dulbecco's modified Eagle's medium plus 10% fetal calf serum supplemented with glutamine in an atmosphere containing 5% CO\(_2\). Cells were transfected using a standard calcium phosphate procedure (21). After 48, transiently transfected cells were harvested, and total RNA was isolated using a guanidinium-based procedure (21). Stable lines were established by incubating cells 48 h post-transfection in 750 \( \mu \)g/ml (effective concentration) Geneticin (Life Technologies, Inc.) and allowing the growth of distinct colonies (approximately 3 weeks). RNA was isolated from pooled colonies using the same guanidinium procedure.

**Qualitative RT-PCR Analysis**

Preparation of total RNA and reverse transcription reaction conditions for the polymerase chain reaction and subsequent gel analysis of PCR products were carried out as described (20). Two sets of primers, both specific for the human AMPD homolog, were used in these studies. PCR products of 69 bp (exon 2-plus) and 57 bp (exon 2-minus) were generated as described (20). RT-PCR analysis yielding products of 216 bp (exon 2-plus) or 204 (exon 2-minus) was generated using the primer H6 (5'-GTCTGGATCTCATCCACATC-3', which bound to exon 3) in the

**Fig. 1. Effect of deletions within intron 2 on exon recognition. Panel A, map of the vector harboring the AMPD mini-gene construct used as the basis for all of the splicing analysis in this paper. Panel B, different fragments of intron 2, as defined by the indicated restriction sites, were deleted alone or in combination from the wild-type mini-gene as illustrated. Fragment 2 was excised and reversed (Rev). Qualitative RT-PCR analysis was performed from RNA isolated from pooled colonies of stably transfected Balb/c 3T3 cells or Soleus 8 myoblasts as described under “Experimental Procedures.” Panel C, radiolabeled PCR products were separated on a nondenaturing polyacrylamide gel and visualized by autoradiography. Restriction enzymes are as follows: RV, EcoRV; RI, EcoRI; B, BamHI; Kp, KpnI.**
RT reaction, and H6 and TH2 (5'-GTCACCCCCACAGTCTCCTC-3') which bound to exon 1) in the PCR. The cycling conditions used with the second set of primers were: 93 °C, 3 min, 1 cycle; 93 °C, 1 min, 58 °C, 1 min, 72 °C, 30 s, 30 cycles; 72 °C, 10 min; otherwise the reagents, buffers, and conditions were identical to those described previously (20).

In either case, these primers were specific to the human homolog of AMPD and did not amplified endogenous murine AMPD.

**Linker Scan Mutagenesis**

The p25 plasmid (see above) was used as the starting vector from which a series of nested deletions, averaging 50 bases, was generated from either end of the AC fragment, using the exonuclease III-based Erase-a-Base kit (Promega). After digesting the single-stranded 5'-overhang with S1 nuclease, double-stranded linkers incorporating an 8-base ASC restriction site (5'-GCTGACGCCCGGCG-3') were ligated onto the resulting blunt ends. After digesting the deletions with ASC I (New England Biolabs) and SacI (which cuts only in the ampicillin resistance gene of the vector) they were sized on agarose gels, and the appropriate deletion pairs were ligated. Using the EcoRI and BamHI sites, each of the resulting 22 ASC mutations replaced the wild-type AC region in the AMPD mini-gene expression vector.

**Site-directed Mutagenesis**

The mutations, designed to alter contiguous bases in groups of four, were introduced in and around the mutation 7 and 8 ASC sites and were generated using the QuikChange site-directed mutagenesis kit (Stratagene). The top strand of each of the 10 pairs of double-stranded oligonucleotides used in this procedure, incorporating the 4 mutated bases (underlined) flanked by 15 wild-type nucleotides, were as follows (the numbers correspond to the number of the mutation in Fig. 6A).

1) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
2) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
3) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
4) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
5) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
6) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
7) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
8) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
9) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'
10) 5'-GAGTCTCTGCTCTGCTTAAATGAGTGCTGGATCCAGT 3'

The bottom strand for each oligonucleotide is simply its complement. All mutations were generated in the p25 construct and verified by sequencing. As in the previous manipulations of this region, each of the 10 mutations was subcloned using the 5'-EcoRI and 3'-BamHI sites flanking the AC region into the AMPD mini-gene expression vector.

**Preparation of Nuclear Extracts**

Confluent cultures of murine Soleus 8 myoblasts were allowed to fuse into myotubes by reducing the fetal calf serum levels in the culture medium from 20% to 2% (23). 72 h after the medium switch, the culture dishes were cooled on ice, and the cells were lysed in ice-cold Triton extraction buffer (T buffer) (20 mM HEPES (pH 7.9), 10 mM NaCl, 3 mM MgCl2, 0.2 mM EDTA, 0.1% Triton X-100, 20% glycerol, 1 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride, 5 μM leupeptin, 5 μg/ml pepstatin, 10 μg/ml aprotinin) with 20 strokes of a Dounce homogenizer (type B pestle). Nuclei were pelleted by centrifuging the lysed cells at 3,000 × g for 15 min at 4 °C. Nuclei were washed once in T buffer, re-pelleted, and then extracted by resuspending in 1 ml of TS.
buffer (T buffer with high salt, 500 mM NaCl and 0.4 mM KCl)/3 × 10^7 cells, in a 15-ml screw-capped tube. After gently rocking the nuclear extraction at 4 °C for 1 h, it was spun at 13,000 × g for 10 min at 4 °C, and the supernatant was aliquoted, snap frozen in liquid nitrogen, and stored at −80 °C. An aliquot of this extract was used only once in subsequent experiments.

Preparation of Radiolabeled Transcripts for Gel Shift Probes

The following oligonucleotides were designed with a T7 promoter at the 5'-end (italicized) followed by 30 bases from intron 2 of the AMPD gene. Centered in these 30 bases were the wild-type sequences defined by the 5'-AATGGAACACCAAGTAA-3' and the mutant sequences defined by the 5'-AATGGAACACCggcg-3'.

Two additional pairs of oligonucleotides were synthesized, identical to the sequences above, except that the 8 bases defining mutation 7 and 18 sites were altered to match their cognate linker scan mutations. The mutated bases are indicated below in lowercase letters.

**Mutant Probe 7**—Top, 5'-GTAATACGACTCACTATAAGGGGCGCAGGC-3'; bottom, 5'-GAGATTATGCTAGTCACTCAGGTGGCAGC-3'.

**Mutant Probe 18**—Top, 5'-GTAATACGACTCACTATAAGGGGCGCAGGC-3'; bottom, 5'-GAATAGAACACCAAGTAA-3'.

**Probe 7**—Top, 5'-GTAATACGACTCACTATAAGGGGCGCAGGC-3'; bottom, 5'-GAGATTATGCTAGTCACTCAGGTGGCAGC-3'.

**Probe 18**—Top, 5'-GTAATACGACTCACTATAAGGGGCGCAGGC-3'; bottom, 5'-GAATAGAACACCAAGTAA-3'.

**A Novel Enhancer of Mini-exon Splicing**

FIG. 3. Distinct regions defined through gross deletions in fragments A and C promote exon 2 inclusion. Panel A, schematic of the deletions made in the fused AC fragment. The size of the A fragment is 591 bases; the C fragment is 402 bases. Deletions were made in either the 5' or 3' direction by, on average, successive 50-base increments. The size of each deletion is indicated by upward arrows in Fig. 4A. The numbers (underlined) in panel B. Each deletion was used, in turn, to replace the wild-type AC fragment in the AMPD mini-gene construct and stably transformed into Balb/c 3T3 cells or Soleus 8 myoblasts. Total RNA from pooled G418-resistant colonies was then analyzed by RT-PCR assay. Panel B, results of the qualitative RT-PCR assay from myoblast RNA reveal that sequences near the middle of the A element (lanes 2 and 3) and toward the 3'-end of the C element (lanes 4 and 5) are each required for the complete inclusion of the 12-base mini-exon 2 in the final splicing product. The results obtained from 3T3 fibroblasts and the myoblasts were identical, nt, nucleotide. of nuclear extract, water to 15 μl, and ~2.5 × 10⁶ cpm of probe was incubated on ice for 20 min and then loaded directly onto a pre-run 5% nondenaturing gel (80 mM Tris-borate, 2 mM EDTA).

RESULTS

Inclusion of Exon 2 Requires a Complex Control Element in the Downstream Intron—Using a mini-gene construct comprising part of exon 1, exon 2, intron 2, and part of exon 3 (Fig. 1A), we had demonstrated previously, in both fibroblasts and myocytes, an absolute requirement for the 5.2-kb second intron for inclusion of exon 2 in the three-exon splicing product (20). Interestingly, as the phenotype of a cell became more muscle-like, the requirement for intron 1 for the inclusion of exon 2 in the final splicing product dropped dramatically; that is, when exon 1 was deleted from the mini-gene construct, 3T3 fibroblasts excluded exon 2 totally, whereas skeletal myoblasts included exon 2 slightly greater than half of the time. When these same myoblasts (Soleus 8 cells) were allowed to differentiate and fuse into myotubes, greater than 90% of final splicing products included the mini-exon (20).

As a first approach at identifying the sequences responsible for this effect, we constructed several mini-gene constructs carrying gross deletions of intron 2 and performed qualitative RT-PCR analysis to test the effects of each on exon 2 inclusion. The primers employed in these assays were specific for the human AMPD gene and did cross-hybridize with the murine homolog (see “Experimental Procedures” and Ref. 20). Nearly half of the sequences of intron 2, those closest to the flanking exons, could be eliminated without an effect on the wild-type ratio of exon 2-included to exon 2-excluded splicing products. Deletions that remove virtually all of the sequences flanking this central 2.7 kb (Del 1/3, Fig. 1B) retain the wild-type splicing pattern in both myoblast and fibroblast cells (Fig. 1C, Del 1/3 lanes). Located in the middle of intron 2, this 2.7-kb fragment must, in addition, be oriented in the wild-type 5′→3′ direction because flipping it 180° caused a total elimination of exon 2 in the final splicing product (Fig. 1, panels B and C, Rev). A myoblast/myocyte-specific enhancement of exon 2 inclusion was observed consistently throughout these studies.

Using convenient restriction sites, the 2.7-kb region was
**Fig. 4.** The positions of two critical linker scan mutations through A and C are consistent with the gross deletion data. Panel A, sequence of the human 989-base AC region showing the position of the 22 linker scan-generated mutations. The downward arrow denotes the junction of the A and C fragments, as defined in Fig. 2; the upward arrows demarcate end points of the 150-bp deletions used in Fig. 3. An alteration in a given base in the linker scan mutations is denoted by the presence of at least part of the ASC restriction enzyme recognition site (GGCGCGCC) above the sequence. TGCATG hexamers are highlighted in boldface and underlined, and marked with an asterisk. Pentamer and 4-base partial repeats of that hexamer sequence are highlighted in boldface and underlined. The two linker scan mutations that, individually, disrupted total inclusion of exon 2 (see panel B), mutation 7 and 18 sites, are boxed. Panel B, results of the qualitative RT-PCR analysis of the linker scan mutations from RNA isolated from pooled colonies of stably transfected 3T3 cells. AC fragments harboring, in turn, each of the individual mutations depicted in panel A were cloned in place of the wild-type AC region in the mini-gene construct and tested for their ability to affect the complete inclusion of exon 2. Only two, mutations 7 and 18, had any appreciable affect, disrupting exon 2 inclusion by greater than 50%. B2 is the mini-gene construct that harbors the entire wild-type sequence of intron 2 depicted in Fig. 1A, giving the partial exon 2 inclusion splicing pattern seen here.

divided into four fragments and analyzed further through their systematic deletion (Fig. 2A). Only when the 600-bp fragment A and 400-bp fragment C were both present in the mini-gene construct did exon 2 appear in the final splicing product in either myoblasts or fibroblasts (Fig. 2). The 800-bp fragment separating A and C, fragment B (the lane labeled ΔΔΔ) cannot, by itself, promote exon 2 inclusion to any extent measured by this assay. Interestingly, in the absence of the B, the presence of A and C together promoted the inclusion of exon 2 to greater than 95% in fibroblasts (over double the rate in the presence of fragment B), suggesting that B may function to attenuate that effect in these cells (see “Discussion”). Taken as a whole, these data indicate that the regulation of exon 2 inclusion in the final splicing product is complex, involving both positive and negative influences. We chose to focus our efforts on narrowing down those sequences in A and C responsible for the enhancement of exon 2 inclusion in the final splicing product.

**Distinct Regions in Fragments A and C Promote Exon 2 Inclusion**—To simplify the identification of the sequences responsible for this effect, we began with the mini-gene splicing substrate in which the 2.7-kb fragment in intron 2 is replaced with the positive acting A and C elements (the ΔB ΔD construct in Fig. 2A). In this context, the combined 1,000 bases (approximately) of sequence promotes virtually total inclusion of exon 2, independent of cell type (see Fig. 2B). Using PCR-generated fragments, we created a series of progressive deletions in A and C in ~150-base increments (the end points of which are demarcated by upward arrows in Fig. 4A). Thus, the 600-base A fragment was subjected to four incremental deletions, starting from its 5′-end, whereas the 400-base C fragment had three, proceeding in the opposite direction (Fig. 3A). These deletions were then used, in turn, to replace the wild-type A and C elements for subsequent splicing analysis. The results in Fig. 3B clearly show that the sequences responsible for the enhancing activity of A reside within the second and third deletion fragments (labeled 2 and 3 in the schematic in Fig. 3A), whereas the positive acting sequences in fragment C lie within the 3′-most 150 bases of that fragment (labeled 5 in Fig. 3A). Systematic Mutagenesis Confirms the Number and Position of Positive Acting Sites—In addition to revealing that the first 150 bases of the A fragment were dispensable for the promotion of exon 2 inclusion, the gross deletion analysis indicated approximately where two critical regulatory sequences in A and C resided. To narrow these further, we performed a saturation mutagenesis on the entire A–C region (minus the 5′-most 150 bases) using a standard linker scan procedure. In the process, we introduced an ASC restriction site (5′-GGCGCGCC-3′) every 38 bases, on average, for a total of 22 different mutations (Fig. 4A; the end points of the 150 base gross deletions are indicated by upward arrows). Each of these was then tested, as before, in the context of the AMPD mini-gene splicing substrate. In only two cases, 7 and 18, was exon 2 excluded to any extent from the final splicing product (Fig. 4B). Both of these mutations had a more subtle effect on exon 2 inclusion than the 150-bp deletions depicted in Fig. 3. The overall results are...
Fig. 5. The critical sites in A and C necessary for exon II inclusion require a minimal distance between them. Panel A, schematic of the constructs derived by removing increasing amounts of sequence between 7 and 18 sites. AC1 and AC5 are controls derived from the 150-base gross deletions depicted in Fig. 3 (called 1 and 5, respectively, in that figure). AC1 gives total exon 2 inclusion in the final mini-gene splicing product. AC5 gives total exclusion. B2 is the mini-gene construct that harbors the entire wild-type sequence of intron 2 depicted in Fig. 1A. The asterisks in panel A mark the relative positions of the ASC linker scan mutations 7 and 18. Numbers to the left of each construct and above the lanes in panel B are the amount of sequence removed from between these two sites. Panel B, results of the qualitative RT-PCR analysis (3T3 cells) using each of the constructs depicted in panel A. When 7 and 18 sites are separated by a distance of between 150 and 110 bases, exon 2 is completely excluded in the final splicing product. wt, wild-type.

consistent between the two approaches. One apparent distinction is that the large deletion of 300 bases at the 5’-end of fragment A (Fig. 3A, construct 2) reduced exon 2 inclusion by about half, whereas the individual linker scan mutations in this same region (Fig. 4, mutations 1–6) produced no effect on total inclusion (see “Discussion”).

A Minimal Functional Distance Is Required between the Two Critical Sites—In the wild-type splicing substrate the sequences surrounding ASC mutations 7 and 18 are separated by roughly 1,150 bases (including the 800-base B fragment (see Fig. 2A)). Because both sites are required for the inclusion of exon 2, we decided to ask directly how critical the spacing between them is. The 600-base A and 400-base C fragments catenated together promote complete inclusion of exon 2 (ΔB ΔD construct, Fig. 2A). Beginning with this construct (in which the 7 and 18 sites are separated by 431 bases (labeled wt AC in Fig. 5A)), we systematically reduced the distance between the centers of ASC mutations 7 and 18 and found that the two sites could be separated by as little as 150 bases and still give complete inclusion of the mini-exon (Fig. 5B).

Site-directed Mutagenesis around 7 and 18 Sites Reveals Their Uniqueness—We then sought to narrow further still the sequences responsible for exon 2 inclusion by generating five 4-base mutations surrounding ASC linker scan mutation 7 and a second set of five mutations surrounding mutation 18. These site-directed mutations were designed so that every base was altered (Fig. 6A). As before, each individual mutation was tested in the context of the mini-gene for its ability to affect the splicing outcome. Except for the 3’-most of these mutations (10 in Fig. 6A), all attenuated the inclusion of the mini-exon 2 to some extent. Four of these (mutations 2, 3, 8, and 9) overlapped with the ASC mutations 7 and 18, corroborating those results. Of the remaining five site-directed mutations, the first and seventh (1 and 7 in Fig. 6, A–C) led to total exclusion. Combin-
3′-acceptor and the 5′-donor splicing sites that flank it (20). The intron-exon boundaries of exon 2 may have evolved so that they are only weakly recognized by the constitutive splicing apparatus, allowing it to respond to the dynamic metabolic requirements of myocytes while preventing its inclusion in inappropriate cell types or physiological circumstances. Reinforcing this notion was our earlier finding that transposing exon 2 into the middle of a different intron of the AMPD gene led to its total exclusion from the final splicing product (17). This result led us to hypothesize that sequences surrounding exon 2 (an ExRE) in the wild-type AMPD gene must act to “unmask” this exon and allow it to be recognized by the splicing apparatus. We went on to show that such recognition-enhancing sequences reside within the 5.2-kb second intron and that their effect on exon 2 inclusion is most pronounced in myocytes (20).

In the present work, we narrowed the ExRE to a roughly 2.7-kb region in the middle of intron 2 and showed that the normal 5′ → 3′ orientation of this sequence is critical (Fig. 1). When this 2.7-kb fragment is replaced with an unrelated sequence of similar size, the mini-exon is again absent from the final splicing product (data not shown). Dividing this 2.7-kb region using convenient restriction sites (Fig. 2, bottom) led to its total exclusion from the final splicing product (17). This 2.7-kb fragment, the A and C fragments, were both necessary and sufficient for the enhancing activity of the ExRE (Fig. 2, B and C).

When brought together outside the context of the 2.7-kb fragment, the A and C fragments promoted total inclusion of exon 2 in fibroblasts (Figs. 2 and 3). A systematic dissection of this 1-kb fragment, first through gross 150-base deletions (Fig. 3) followed by linker scanning mutagenesis (Fig. 4), narrowed the functional sequences down to two distinct, widely spaced sites, one in A (mutation 7 from the linker scan, Fig. 4A), the other in C (mutation 18). The locations of these two sites are consistent with those regions from the gross deletion analysis shown to be required for exon 2 inclusion. One apparent exception is that removal of the 5′-most 300 bases of the A fragment reduced the ability of the splicing machinery to recognize exon 2 (Fig. 3B, construct 2), whereas individual linker scan mutations 1–6 (Fig. 4) in this same region had no effect. We hypothesize that either the linker scan mutations we generated in this instance left critical sequences intact, or the influence of each individually on exon 2 recognition is redundant. These sites and the sequences immediately surrounding them were then dissected further via site-directed mutagenesis. Only one of these mutants, centered 15 bases downstream of the 18 site (mutation 10), had no effect on the total inclusion of exon 2 (Fig. 6, and see below).

The wild-type intron contains the nearly 800-ase B element interposed between the positive acting A and C fragments (Fig. 2). When we moved the mutation 7 and 18 sites closer together by removing the B element, recognition and inclusion of exon 2 in fibroblasts reached the levels observed in Soleus 8 myoblasts, to 100% of the splice products (Fig. 2A, ΔB construct and ΔB ΔD construct). Total inclusion continued to be observed as the two sites were moved progressively closer together. At the transition from 150 to 110 bases the enhancing activity is lost completely. Because both sites are required for exon 2 inclusion, one possibility, based on the spacing data, is that torsional constraints, imposed by RNA secondary structure, prevent a required interaction between them below a separation of about 150 nucleotides. Looping of the RNA in this region may bring together two positive acting sites that, individually, cannot promote exon 2 inclusion (11).

Cis-acting enhancing elements that direct the splicing machinery to proximal, usually suboptimal, splice sites fall into two general groups, exonic and intronic. Most exonic enhancers are comprised of purine-rich sequences that are thought to mediate their activity mainly through the binding of constitutive SR proteins (11, 30–32). A number of intronic splicing enhancers from both vertebrates and invertebrates which influence alternative splicing in a cell-specific manner have been described (8, 33–39). In the case of the alternative splicing of very small exons, several, including AMPD, are known to be regulated by downstream intronic enhancers (6, 7, 40–42). In the majority of these, two recurring types of sequence elements mediate alternative splicing. The first is a polypyrimidine tract (37, 43) and short sequence repeats (38, 39, 41, 42, 44, 45, 46). There is a polypyrimidine tract within the A element described here (mainly T₃, between the linker scan mutations 2 and 7, see Fig. 4A), but mutations within this region did not produce an effect on exon 2 inclusion (Fig. 4B).

Another recurring downstream intronic enhancing element is the hexamer TGCATG. This sequence has been demonstrated to function as an enhancer in the alternative splicing of the neuron-specific 3-base N1 exon of the mouse c-src gene (47),
the 30-base N30 exon of the human non-muscle myosin heavy chain gene (48), and exon IIIB of the fibronectin gene (38, 39, 45, 49). The presence of one or more copies of a TGCATG enhancing site in the intron immediately downstream of exon IIIB of the fibronectin gene is conserved phylogenetically from frog to human (45). Like SR-binding elements, the TGCATG hexamer and its variants may, in certain contexts, only function in multiple copies to attract a threshold number of factors to initiate a splicing event (9). In the nearly 1,000 bases that comprise the A and C fragments from the AMPD intronic enhancer, the TGCATG hexamer appears twice, once about 100 bases upstream of the linker scan mutation 18 and again about 100 bases downstream of that same site (both in the C fragment; see the asterisks in Fig. 4A). Linker scan mutation 15 overlaps (by 2 bases) the first hexamer, whereas mutation 22 is centered 17 bases upstream of the second hexamer. Neither of these mutations, on its own, affected the splicing outcome (Fig. 4).

Partial repeats of the hexamer sequence, GCATG (once), TGCAT (twice), and TCGA (four times) appear throughout this kilobase of sequence. One 4-base TCGA sequence is centered at the mutation 7 site, and the two 3′-most bases of TCGA overlap the 18 site (Fig. 4A). However, 9 out of 10 of the site-directed mutations introduced around the mutation 7 and 18 sites disrupt the enhancing activity of the AC element at least as well as these two linker scan mutations. This result argues against the TCGA and TGCAT partial hexamer sites alone having sufficient functional enhancing activity, although disruption of the sequences flanking them may be enough to abrogate such an activity.

Sequences surrounding the 7 and the 18 sites are both required for the inclusion, to any extent, of the mini-exon 2. The requirement for both sites in the detection of a binding activity from muscle cells is consistent with all of the previous results we have observed in the detection of the AMPD mini-gene splicing products. It remains to be proved, of course, whether the binding activity we have detected here actually contributes to the enhancing effect.

Our survey of the literature has produced no published data that identify the TGCATG sequence variants TGCAT or TGCA specifically as binding sites for factors involved or potentially involved in the regulation of splicing. Nor is there any other significant sequence similarity between the 7 and 18 sites. Furthermore, we do not detect any occurrence of the sequences surrounding the 7 or 18 sites in any other published splicing enhancer. Additional mutational analysis should reveal whether these short sequence variants are functionally responsible for the enhancing effect on exon 2 inclusion. The requirement for both of these sites to achieve a specific binding activity provides additional compelling evidence that they in some way interact with one another, perhaps through the binding of one or more proteins (or ribonucleoproteins).

A bipartite splicing enhancer element has been identified in the intron downstream of exon IIIB of the fibroblast growth factor receptor gene (50). In this case, mutation of either one of
the two sites alone diminishes, but does not eliminate, the activity of this enhancing element. This is in contrast to the data reported here in which the disruption or deletion of sequences surrounding mutation 7 or 18 abolishes the enhancing activity completely. We have demonstrated that both sites, normally separated by about 1,150 bases, are required for intron 2 inclusion in all of our splicing assays as well as for the detection of an RNA binding activity. To our knowledge, however, this is the first report of a bipartite splicing enhancer sequence where both of the unique sequences are required to obtain a binding activity.

It is possible that in vivo, non-muscle cells can tolerate only minimal levels of AMPD exon 2-containing protein (19). To ensure then that the inclusion of exon 2 was less efficient, the AMPD gene may have evolved so that the enhancing activity of the A and C elements works efficiently only in the context of muscle and then only under certain physiological conditions. In any case, it is clear that cis-regulation of the alternative splicing of exon 2 is complex and likely involves more elements, both positive and negative, than have been presented evidence here that the ultimate inclusion of this final splicing product of our mini-gene construct (20). We have obtained a binding activity.

It is possible that in RNA binding activity. To our knowledge, how-