Regulation of Alternative Splicing: More than Just the ABCs*

Published, JBC Papers in Press, November 16, 2007, DOI 10.1074/jbc.R700031200
Amy E. House and Kristen W. Lynch 1
From the Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, Texas 75390-9038

Alternative pre-mRNA splicing, the differential inclusion or exclusion of portions of a nascent transcript into the final protein-coding mRNA, is widely recognized to be a ubiquitous mechanism for controlling protein expression. Thus, understanding the molecular basis of alternative splicing is essential for deciphering post-transcriptional control of the genome. Pre-mRNA splicing in general is catalyzed by a large dynamic macromolecular machine known as the spliceosome. Notably, the recognition of the intron substrate by spliceosomal components and the assembly of these components to form a catalytic spliceosome occur through a network of highly combinatorial molecular interactions. Many, if not all, of these interactions are subject to regulation, forming the basis of alternative splicing. This minireview focuses on recent advances in our understanding of the diversity of mechanisms by which the spliceosome can be regulated so as to achieve precise control of alternative splicing under a range of cellular conditions.

Pre-mRNA splicing is a critical step in the expression of nearly all eukaryotic genes in which intron sequences are removed and exons are joined together to generate a mature protein-coding mRNA transcript. The chemistry of the splicing reaction is mediated by the “spliceosome,” an RNA-based molecular machine containing five snRNAs and numerous associated proteins (1). Both the snRNA and protein components of the spliceosome interact with defined sequences at the exon/intron boundaries to direct RNA excision and ligation at these “splice sites” (Fig. 1a). In addition, several of the snRNAs interact with one another to ensure the correct juxtaposition of distant regions of the substrate required for splicing catalysis.

Although the spliceosome catalyzes RNA cleavage and ligation with high fidelity, the inherent flexibility of this enzymatic complex allows it to be highly sensitive to regulation (2). A frequent consequence of spliceosome regulation is the differential inclusion or exclusion of exons in the final mRNA product in a process known as alternative splicing. Alternative splicing is predicted to occur in the vast majority of mammalian genes and is a primary mechanism by which complex organisms can regulate protein expression and generate a diverse proteome from a relatively limited genome (2). Although initial studies of alternative splicing suggested that regulation occurred predominantly at the earliest steps of spliceosome assembly, more recent studies have demonstrated regulation of splicing patterns at many points throughout the assembly pathway. In this minireview, we will walk through the spliceosome assembly pathway to highlight both “traditional” and newly appreciated mechanisms of alternative splicing, and we will discuss what recent advances in our knowledge of transitions in the general spliceosome assembly pathway reveal about the potential for even further mechanisms of splicing regulation.

Dynamic Assembly of the Spliceosome

Each of the snRNAs that compose the spliceosome (U1, U2, U4, U5, and U6 snRNAs) associates with a number of proteins to form a ribonucleoparticle called an “snRNP.” The catalytic conformation of the spliceosome (so-called “C” complex) does not exist de novo in its final structure, but rather forms in a highly dynamic process best described by a stepwise pathway involving several intermediate complexes (E-A-B) that have been identified and characterized in vitro and in vivo (Fig. 1b) (2, 3). The earliest known complex committed to the splicing pathway (E) is defined by U1 snRNP base-paired to a 5′-splice site, with the 3′-splice site recognized by binding of the U2AF heterodimer (U2AF65/35) to the polypyrimidine tract and 3′-terminal AG, respectively, and association of the protein SF1 with the branch-point sequence (BPS). The E complex is chased into the pre-spliceosome A complex by the ATP-dependent addition of U2 snRNP at the 3′-splice site facilitated by base pairing between the U2 snRNA and BPS. Recruitment and addition of the U4-U6/U5 tri-snRNP, which contains the remaining spliceosome subunits, results in formation of the B complex. Finally, the C complex forms by extensive remodeling of both the snRNA and the protein components that are present in the B complex, including loss of both the U4 and U1 snRNPs, to produce an active site that is capable of catalyzing the transesterification chemistry required for exon ligation and lariat release. The release of U1 and U4 snRNPs, as well as many other molecular rearrangements required for assembly, is promoted by the action of a series of DEX(D/H) box ATPase proteins, which will be discussed further below (4, 5).

By definition, the complexes described above represent a spliceosome formed around an intron in an orientation described as “intron-defined.” This definition is based on the early use of genes from Saccharomyces cerevisiae or artificial metazoan model substrates (which typically contain a single short intron) in studies of spliceosome assembly. However, the subsequent use of more complex splicing substrates has led to the conclusion that, at least during the earliest steps in assembly, the metazoan spliceosome is built around the exon in a manner termed “exon definition” (Fig. 1b) (6, 7). The final C complex must be formed around the intron for proper catalysis; however, how long the exon-defined conformation persists...
MINIREVIEW: Regulation of Alternative Splicing

![Diagram of spliceosome assembly](attachment:spliceosome_diagram.png)

**FIGURE 1. Basic exon recognition and spliceosome assembly.**

**a**. Consensus sequences that define exon/intron boundaries. Exons are noted by boxes, and introns by lines. Consensus nucleotides are indicated above the line (Y = U or C, R = G or A), and the term for sequence is shown below. ss, splice site; PPT, polypyrimidine tract. **b**, schematic of the steps described in spliceosome assembly. Left, canonical intron-defined orientation; right, corresponding exon-defined version of each step. For simplicity, only U1, U2, U4, U5, and U6 snRNPs and U2AF are shown. **c**, examples of the basic function of SR proteins bound to ESEs (green) and hnRNPs bound to ESSs (red). SR proteins promote exon definition by recruiting U2AF and U1 via protein/protein interactions and U2 snRNPs via protein/RNA interactions. hnRNPs can inhibit exon definition by sterically blocking SR protein or U2AF interaction with the substrate.

During spliceosome assembly remains an open question. Although kinetic studies have suggested that the commitment of splice site pairing occurs at the A complex stage in assembly (8), it has also been demonstrated that U1 and U2 snRNPs can bind the 5′- and 3′-splice sites flanking an isolated exon to form a stable complex that is subsequently capable of splicing in trans to a separate exon (9). It is possible that the stage at which the growing spliceosome transitions from exon definition to intron definition differs between substrates in a manner determined by factors such as intron length, auxiliary regulatory proteins, and splice site strength. Moreover, cross-exon and cross-intron interactions between snRNPs components also may not be mutually exclusive, but rather may occur simultaneously via distinct faces of the snRNPs to assist in the overall assembly of the spliceosome. Finally, it is worth noting that the complexes shown in Fig. 1b are unlikely a comprehensive description of spliceosome assembly. As our ability to isolate and characterize the spliceosome increases, so does our appreciation of previously unidentified transition states in the assembly pathway (5). Since each molecular rearrangement and transition during spliceosome assembly represents a potential point of regulation, a more detailed characterization of spliceosome assembly will ultimately lead to a deeper understanding of the mechanisms of alternative splicing.

**Auxiliary Sequences and Proteins in Alternative Splicing**

Although the splice sites within the pre-mRNA function to direct the splicing machinery, these sequence elements in higher eukaryotes are highly degenerate and often imbedded within introns that are significantly longer than exons. Thus, frequently as few as a handful of nucleotides mark the ends of an intron often tens of thousands of bases long (2). Therefore, it is not surprising that sequence elements outside of the splice sites can strongly affect metazoan pre-mRNA splicing. *cis*-Acting auxiliary sequences occur within both exonic and intronic regions and can either promote recruitment of the spliceosome and exon inclusion (splicing enhancers) or disrupt assembly of the splice machinery and cause exon skipping (splicing silencers). Use of most exons is now believed to be under the combinatorial control of multiple regulatory RNA elements as well as the inherent strength or weakness of the flanking splice sites (6, 10).

Although a few regulatory sequences have been shown to function by directly creating RNA secondary structures that alter splice site recognition (6, 11–13), the majority act primarily as platforms for binding of non-snRNP regulatory proteins. To a first approximation, the best characterized of the regulatory elements, exonic splicing enhancers (ESEs), bind a family of proteins known as SR proteins, which contain an RNA-binding domain and a region rich in Arg-Ser dipeptides (RS domain) (2). By contrast, exonic splicing silencers (ESSs) typically function to repress exon inclusion by recruiting members of the hnRNP family of proteins, a structurally diverse set of RNA-binding proteins (2). Although SR proteins and hnRNPs do not always correlate strictly with enhancers and silencers, respectively, this simplification helps illustrate the important emerging concept of a splicing “code” in which the splicing pattern of a gene is determined by the interplay of proteins along a nascent transcript (6, 10). Additional splicing regulatory proteins have also been identified that have similar activity to SR proteins and hnRNPs but do not fall cleanly into one of these two protein families (2). Such non-SR/hnRNP splicing regulatory proteins further increase the complexity of the splicing regulatory machinery.

**Regulation of Splice Site Recognition**

The first and best characterized splicing enhancers and silencers are those that control the earliest steps of spliceosome assembly, *viz.* the association of the U1 snRNP, U2AF, and the U2 snRNP with the 5′- and 3′-splice sites, respectively (2, 10). SR proteins have been shown to interact with both the U1 snRNP and the U2AF heterodimer, thus recruiting these splicingomalous components to a particular exon (Fig. 1c) (2). Furthermore, the RS domains of SR proteins stabilize RNA base pairing interactions between the U2 snRNA and BPS (14). SR proteins are thought to bind to most exons to promote basic exon definition interactions even for constitutive (non-alternative) exons (16). However, the association of SR proteins with
ESEs is often relatively weak and thus can be promoted or blocked by neighboring proteins to regulate exon inclusion. For example, the Drosophila female-specific splicing regulatory protein TRA stabilizes the binding of SR proteins to an ESE to facilitate recruitment of the U2AF heterodimer to the weak female-specific polypyrimidine tract of the Doublesex gene (2). By contrast, inclusion of human immunodeficiency virus-1 tat exon 2 is inhibited by hnRNP A1 competing with the SR protein SC35 for binding to an overlapping ESE/ESS element, thereby preventing SR protein-dependent recruitment of U2AF to a weak 3′-splice site (15). These examples also highlight the combinatorial control of splicing regulation by cooperative or dual cis-acting elements. An implication of this combinatorial interpretation of the splicing code is that the precise balance of regulatory proteins in any given cell can have a profound influence on the ultimate splicing pattern of a gene.

In addition to the ability of hnRNPs to compete directly with SR proteins, there have been several well studied examples of hnRNPs functioning directly to repress U1 or U2 binding to an exon (Fig. 1c) (2). In the simplest model, binding of an hnRNP to an ESS, or an intronic splicing silencer located close to the exon, causes a direct steric block in the ability of a spliceosomal component to bind to an overlapping sequence, similar in concept to the competition between hnRNPs and SR proteins described above. For example, binding of hnRNP H to the extreme 3′-end of NF-1 exon 3 blocks U1 binding to the adjacent 5′-splice site (16). Oligomerization of hnRNPs along the pre-mRNA can further affect spliceosomal binding to sites distal to the primary location of hnRNP association (17). Alternatively, hnRNPs bound to distant sequences can “loop out” the intervening sequence, as is observed in the autoregulation of hnRNP A1, in which A1 molecules bound to the introns flanking variable exon 7B interact across the exon to sequester it from the rest of the pre-mRNA transcript (18, 19). It should be noted, however, that whether such looping blocks initial access to the splice sites by the snRNPs or prevents appropriate pairing between snRNPs (see below) remains an open question.

**Regulation of Pre-spliceosomal Transition States and Molecular Rearrangements during Assembly**

The examples of regulation outlined above and other similar studies initially led to the general belief that the vast majority of splicing regulation occurs during E or A complex assembly. However, as discussed both above and below, spliceosome assembly is highly dynamic throughout the entire substrate recognition and catalytic cycle. Therefore, it seems likely that many or all of the interactions that are formed and broken throughout assembly are potential points of regulation. Indeed, a growing body of work has now demonstrated regulation of alternative splicing at several points in assembly downstream of the ATP-dependent binding of U2 to the BPS.

Variable exon 4 of the CD45 gene contains a silencer element (known as ESS1) that, when bound by hnRNP L, results in exon skipping (20). Interestingly, binding of hnRNP L to ESS1 does not block either U1 or U2 association with the splice sites flanking the variable exon. Instead, hnRNP L functions to repress exon 4 splicing by causing the formation of a U1-, U2-, and ATP-dependent exon recognition complex that is required to inhibit progression to the U4-U6/U5 tri-snRNP-containing B complex (21). The simplest model for the activity of hnRNP L in the repression of CD45 exon 4 is that this protein interacts with the adjacent U1 and U2 snRNPs, holding them in a conformation across the exon that inhibits cross-intron pairing interactions and/or interactions with spliceosomal components that are required for tri-snRNP recruitment (Fig. 2a). Such a model is consistent with the observed location dependence of hnRNP L function. Although hnRNP L represses splicing when bound to ESS1 of CD45, binding of hnRNP L to a CA-rich enhancer in a central location within an intron promotes excision of the intron and ligation of the flanking exons (22). In this intronic context, co-association of U1, hnRNP L, and U2 would be predicted to promote cross-intron pairing of the U1 and U2 snRNPs, thereby generating a canonical A complex that proceeds efficiently to the catalytic complex.

Enhancement of snRNP pairing has also been proposed as a mechanism for activation of splicing by hnRNP A1 or H when positioned at distal sites within a long intron. In this case, hnRNPs A1 and H are not predicted to directly contact the snRNPs, but rather to dimerize and loop out intervening intron sequences, bringing together the snRNPs bound to the 5′- and 3′-splice sites (23). Therefore, the looping out of a variable exon may not only decrease use of the “sequestered” exon as discussed above, but also promote the alternative pairing of the flanking exons (Fig. 2b). Furthermore, looping out of an exon by flanking hnRNPs is unlikely to simply block access of snRNPs for the repressed exon because in both the autoregulation of hnRNP A1 and in repression of the N1 exon of n-src by the hnRNP PTB to the flanking introns, association of U1 snRNP with the repressed exon is not inhibited (18, 24). Interestingly, detailed analysis of the regulation of n-src demonstrates that intrinsic binding by PTB blocks pairing of the repressed N1 exon with the downstream exon, even after formation of a U1- and U2-containing exon definition complex (Fig. 2b) (24). Thus, exon pairing can be regulated by proteins bound within the intron as well as the exon and may represent a more common mechanism for alternative splicing than previously recognized.

In addition to exon pairing, appropriate recruitment of tri-snRNP is necessary for progression through spliceosome assembly and is also susceptible to regulation. In addition to their role early in assembly, SR proteins have been shown to promote U6 snRNA association with the 5′-splice site and can influence spliceosome formation at this later stage in assembly (14, 25, 26). In contrast, NRS (negative regulator of splicing) within the intron as well as the exon may represent a more common mechanism for alternative splicing than previously recognized.

In addition to exon pairing, appropriate recruitment of tri-snRNP is necessary for progression through spliceosome assembly and is also susceptible to regulation. In addition to their role early in assembly, SR proteins have been shown to promote U6 snRNA association with the 5′-splice site and can influence spliceosome formation at this later stage in assembly (14, 25, 26). In contrast, NRS (negative regulator of splicing) within the intron as well as the exon may represent a more common mechanism for alternative splicing than previously recognized.

In addition to exon pairing, appropriate recruitment of tri-snRNPs is necessary for progression through spliceosome assembly and is also susceptible to regulation. In addition to their role early in assembly, SR proteins have been shown to promote U6 snRNA association with the 5′-splice site and can influence spliceosome formation at this later stage in assembly (14, 25, 26). In contrast, NRS (negative regulator of splicing) within the intron as well as the exon may represent a more common mechanism for alternative splicing than previously recognized.
MINIREVIEW: Regulation of Alternative Splicing

Regulation of Alternative Splicing at Later Steps in Assembly

The regulation of alternative splicing at later steps in assembly is a common regulatory event in spliceosome assembly. By contrast, an increase in the rate of assembly might promote use of an otherwise weak splice site. Such a kinetic model for regulation is similar to a proposed kinetic proofreading model for general splicing (5, 31, 32) and predicts that altering the efficiency at any step during spliceosome assembly could result in a change in splice site choice. Indeed, recent work from the laboratory of T. Nilsen has characterized several splicing silencers that only marginally weaken the inherent efficiency of a neighboring 5’-splice site, but dramatically shift the splicing pattern to an alternate 5’-splice site when one is present. Such studies provide direct evidence for kinetic competition in establishing alternative splicing patterns.

Interestingly, several studies have demonstrated the importance of DEX(D/H) box ATPases in regulating the kinetics of many RNA rearrangements within the spliceosome (31), suggesting that the regulation of these proteins may also play a role in alternative splicing. Elegant genetic work by two groups has shown that toggling between two mutually exclusive structures of the U2 snRNA (stem-loop structures IIa and IIc) promotes distinct steps in both the assembly and catalytic phases of the splicing cycle (33, 34). The DEX(D/H) box ATPases Prp5p and Prp16p, as well as the U2 snRNP Cus2p, have been implicated in regulating these U2 snRNA rearrangements and the progression of spliceosome assembly (33, 34). Similarly, the DEX(D/H) box protein Prp28p is required for the release of U1 snRNA and exchange for U6 snRNA at the 5’-splice site, an activity that is particularly important in cases in which the 5’-splice site deviates significantly from the consensus (35). Finally, the GTPase Snu14p and the associated DEX(D/H) protein Brr2p are required for the dissociation of U4 and U6 that is necessary during the transition to the C complex (36). Interestingly, Snu14p activity in promoting spliceosome assembly is clearly susceptible to regulation by control of its GTP- versus GDP-binding state. More broadly, one can easily imagine that substrate-bound proteins could alter the local recruitment or activity of any of the “checkpoint” DEX(D/H) box proteins mentioned above, thereby increasing or decreasing spliceosome assembly on a nearby exon to regulate its inclusion in the final mRNA. Indeed, a recent knockdown of several DEX(D/H) proteins in Drosophila reveals specific changes in the alternative splicing of some, but not all, variable exons tested (37), providing further evidence for a kinetic model of regulation and for the role of DEX(D/H) proteins in this process.

Regulation of Splice Site Choice during Catalysis

In addition to altering pre-spliceosome formation, regulation of the spliceosome by DEX(D/H) box proteins also occurs during catalysis (31). Although it might not seem possible or prudent to alter splice site choice once catalysis is underway, such regulation has indeed been observed. Strikingly, although binding of U2AF35 to the 3’-AG of an intron is typically required for efficient exon definition, several studies have shown that the identity of the 3’-AG for exon ligation is not irreversibly determined until the actual second catalytic step (38, 39). In other words, the 3’-AG bound early in spliceosome assembly by U2AF35 is not necessarily the same dinucleotide at which the intron is cleaved from the downstream exon.

One of the determinants of the eventual site of second step cleavage is the spliceosomal protein SPF45, which binds to the 3’-AG that is ultimately used for catalysis. In the Drosophila Sex-lethal (Sxl) transcript, SPF45 binds to a 3’-AG upstream from that initially bound by U2AF35 to direct splicing to this proximal site (40). In the absence of SPF45, splicing switches to the downstream 3’-AG with no significant loss of efficiency; however, when the SXL protein interacts with SPF45 at the proximal 3’-AG, it results in complete inhibition of the second step of splicing in this region and in exon skipping (40). Presumably, the association of SXL with SPF45 stalls catalysis with

---

4 T. Nilsen, personal communication.
sufficient local efficiency that eventually a rearrangement occurs to bring the 3'-splice site upstream of the next exon into the catalytic pocket of the spliceosome (Fig. 2c), again consistent with a kinetic model of splicing regulation. It is also worth noting that a rearrangement of SR protein contacts with RNA in the catalytic core of the spliceosome has recently been reported (25), suggesting the possibility that binding of an SR protein to an ESE may, in some contexts, modulate splicing catalysis in addition to exon definition.

Implications of Mechanistic Diversity and Combinatorial Control

Taken together, the studies described above illustrate the diverse mechanisms by which spliceosome assembly and activity can be modulated to achieve differential splicing patterns in a given gene. The examples presented here are not meant to be exhaustive, but rather to highlight common themes and suggest the potential for regulation at any number of as yet unidentified transitions required for building a spliceosome. We also note that, although beyond the scope of this review, additional issues such as speed of transcriptional elongation and co-transcriptional recruitment of splicing factors can also influence the kinetics of spliceosome assembly and splice site choice (41).

Given such mind-boggling diversity of regulatory mechanisms, one might wonder why nature has bothered to find so many routes to the same goal of controlling exon inclusion. One possibility is that the same mechanism for the regulation of any particular gene is dependent on the specific rate-limiting step in spliceosome assembly and catalysis for that gene. Recent global analyses of splicing have shown that knockdown of proteins thought to be core components of the spliceosome results in the expected large defects in splicing of some genes, but has little to no effect on the splicing of other substrates (37, 42, 43). The interpretation of these results is that pre-mRNA substrates differ in their requirement for even central spliceosomal proteins due to redundancy of splicing signals or differences in affinity of substrate/spliceosome interactions. Therefore, by analogy, it is predicted that exons differ widely in their susceptibility to modulation of exon definition, exon pairing, or catalysis. In particular, those exons that need to be included under most conditions are likely to have particularly strong splice sites (i.e. high affinity for U2AF or snRNPs), such that only “later” steps in assembly are available for regulation. Moreover, exon sequences typically have to conform to coding constraints, whereas introns often harbor transcription regulatory sequences or small noncoding RNAs. Thus, sequences that determine alternative splicing regulation are not unlimited in their location and identity, but rather have to accommodate other evolutionary constraints. In essence, therefore, the splicing process can be viewed as a series of checkpoints that allow the spliceosome to sample different choices and to adjust splicing decisions according to local dictates and cellular requirements. Such flexibility and control ultimately provide for the extent of alternative splicing that is now recognized to be pervasive in higher eukaryotes and essential for the functional diversity required of complex organisms.

REFERENCES