Chapter 2
Alternative Splicing of Pre-messenger RNA

Vanessa Cristina Arfelli and Leticia Fröhlich Archangelo

2.1 Splicing and the Splicing Machinery

Since the discovery that eukaryotic genes are discontinuous, much has been learnt about how its transcriptional products are processed to generate the protein-coding mRNAs (Sharp 1994, 2005). Splicing is the cellular process in eukaryotic cells in which the intronic non-coding sequences are removed from the precursor mRNA (pre-mRNA) and exonic sequences are juxtaposed to yield mature functional mRNAs. This process is accomplished by a large machinery, the spliceosome, comprised of five small nuclear ribonucleoprotein particles (U1, U2, U3, U4, and U6 snRNP) and approximately 170 associated proteins, which combinational composition varies from one stage to the next throughout spliceosome cycle (Wahl et al. 2009).

The spliceosome assembly onto the pre-RNA is a highly dynamic process, where a series of consecutive steps produce the complexes E, A, B, C, P and ILS, respectively. The formation of these complexes is based on the establishment and dismantlement of several weak interactions between RNA:RNA, protein:RNA, and protein:protein molecules that act synergically to recognize and assemble onto the pre-RNA splice sites and to form the catalytically active structure (Wahl et al. 2009). The entire process is highly orchestrated and subject to different levels of regulation to guarantee the correct processing of mRNAs and the fidelity of the cellular transcriptome.

Three common consensus sequences define intronic regions of a pre-mRNA and are needed for the initial recognition and assembly by the spliceosome on the splice sites: the 5′ donor splice site (5′-ss) and the 3′ acceptor splice site (3′-ss) at both ends of the intron and a branchpoint sequence (BPS). Every intron almost invariantly contains GU dinucleotide at the 5′ end and AG dinucleotide at the 3′ end. These

V. C. Arfelli · L. F. Archangelo (✉)
Department of Cellular and Molecular Biology and Pathogenic Bioagents, Ribeirão Preto Medical School, University of São Paulo (FMRP-USP), Ribeirão Preto, SP, Brazil
e-mail: leticiafa@fmrp.usp.br

G. A. Passos (ed.), *Transcriptomics in Health and Disease*,
https://doi.org/10.1007/978-3-030-87821-4_2

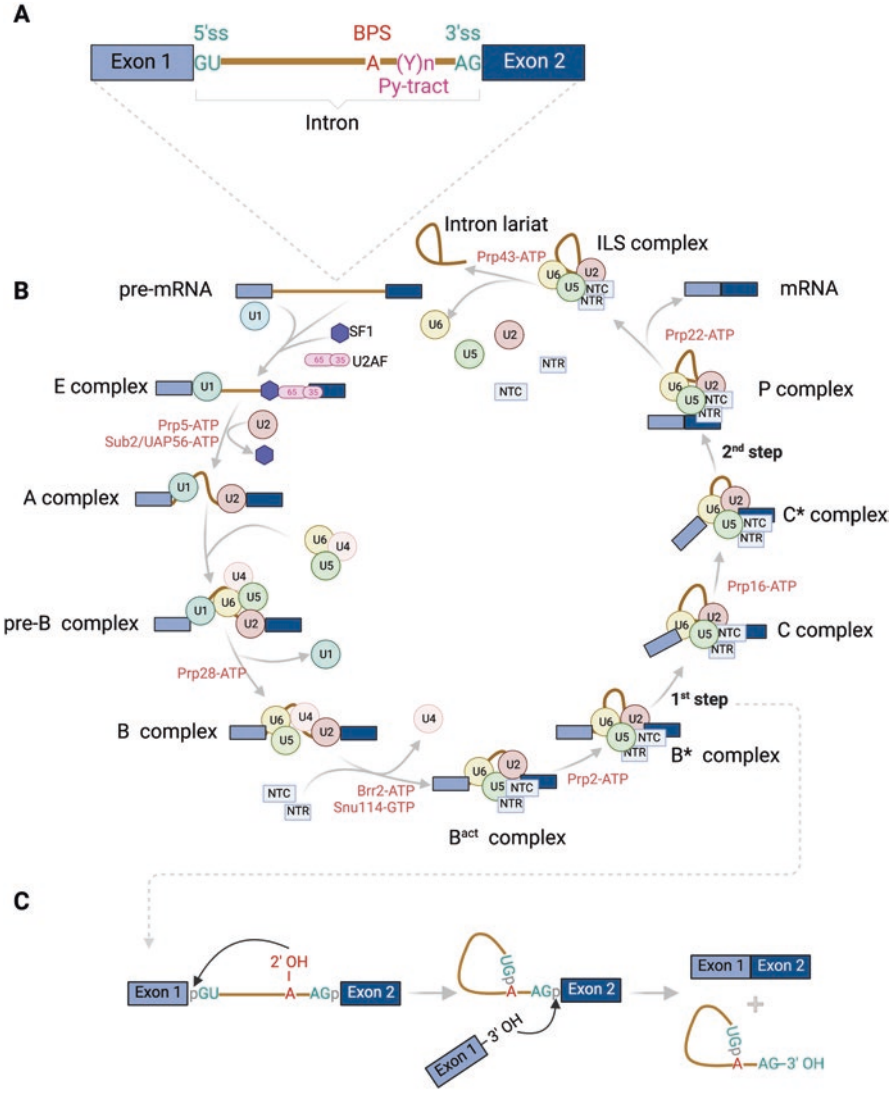

Fig. 2.1 *The splicing process*. (**a**) pre-mRNA structure: Exons 1 and 2 are represented in light and dark blue, respectively. The intron is represented by brown line. The 5′ donor splice site (5′-ss) is defined by the nucleotides GU, while the 3′ acceptor splice site (3′-ss) is defined by the nucleotides AG. In humans, the conserved branch point sequence (BPS) is yUnAy, which is represented in the scheme by the adenine (A). The poly-pyrimidine tract (Py-tract) is represented by a Y. (**b**) Splicing events coordinated by the major spliceosome: the stepwise binding of U2AF (pink), SF1 (purple), snRNPs (colored circles), and NTR and NTC (light blue) is depicted, as well as the action of helicases (red) to form the complexes E, A, B, C, P, and ILS, resulting in the mature mRNA and the removal of introns. (**c**) Transesterification reactions. The adenine (A) 2-OH in BPS mediates the nucleophilic attack to the phosphate group (p) in the 5′-ss of the intron. In the second transesterification reaction, the free 3′-OH of the first exon attacks the phosphate group in the 3′-ss, yielding junction of exons 1 and 2 and removal of the intron in a lariat form. (Made in ©BioRender – biorender.com)

dinucleotides are positioned inside longer consensus sequences known to influence the strength of the splicing sites. The BPS is located in the proximity upstream of the 3′-ss. A pyrimidine enriched sequence, known as polypyrimidine tract (Py-tract), lies between the BPS and the 3′-ss (Fig. 2.1a) (Reed 1996).

The stepwise assembly of the spliceosome begins with the Early complex (complex E) formation when U1 snRNP, through RNA–RNA base-pairing interactions, recognizes the 5′-ss. It is followed by ligation of splicing factor 1 (SF1/BBP) to BPS and its interaction with the U2 auxiliary factor large subunit (U2AF65), which in turn associates with the Py-tract, whereas the small subunit of the U2AF heterodimer (U2AF35) binds to the 3′-ss.

In the subsequent ATP-dependent step, the U2 snRNP is recruited to replace SF1 from its interaction with the BPS. It interacts with U1 snRNP and turns the E complex into a pre-spliceosome complex (complex A). The next step involves binding of the pre-assembled U4/U6 and U5 snRNPs to form the pre-B complex. The exit of U1 snRNP marks the formation of B complex. Although all the snRNPs are present at this point, the complex B is catalytically inactive. In order to activate the spliceosome, the complex has to go through a series of conformational and compositional changes turning it into an activated complex (complex B^{act} and further B*). Specifically, 5′-ss and 3′-ss are brought into proximity, U4/U6 duplex unwind, and the U4 snRNP is released, allowing U6 to interact with the 5′-ss.

Moreover, the NineTeen complex (NTC) and the NTC-related complex (NTR) are recruited. The activated B complex engages in the first catalytic reaction, additional rearrangements occur and generate the catalytic complex C (C* complex), which undergoes the second catalytic step of splicing. In the P complex, interactions between the three conserved elements of the intron (3′-ss, 5′-ss, and lariat junction) occur. Rearrangements promote mature mRNA release, leaving only the intron lariat spliceosome (ILS). Finally, the U2, U5, U6 snRNPs and NTC and NTR complexes are released to engage in an additional round of splicing and the post-spliceosome complex disassemble (Fig. 2.1b) (Matera and Wang 2014; Wahl et al. 2009; Wan et al. 2019).

A large amount of energy is devoted to RNA remodeling throughout spliceosome formation, which is employed by the action of numerous evolutionarily conserved DExD/H type RNA-dependent ATPases/helicases that act at specific steps of the splicing cycle to catalyze RNA–RNA rearrangements and RNP remodeling events (Cordin and Beggs 2013; Staley and Guthrie 1998).

Essentially, the splicing process of intron removal and ligation of the flanking exons entails two trans-esterification reactions involving functional groups in the 5′-ss, 3′-ss, and BPS regions of an intron. First, a nucleophilic attack by the 2′ hydroxyl group of a conserved adenosine residue within the BPS cleaves the phosphodiester bond within the 5′ exon–intron junction. The reaction generates a free 3′ hydroxyl group on the 5′ exon and a lariat intron intermediate. In the second reaction, the phosphodiester bond in the 3′ intron–exon junction is attacked by the 3′ hydroxyl group of the 5′ exon, displacing the lariat and promoting exons ligation (Padgett et al. 1986) (Fig. 2.1c).

The major spliceosome, built by the U1, U2, U4/U6, and U5 snRNPs particles, as described above, is responsible for removing the vast majority of pre-mRNA introns.

However, a distinct but structurally and functionally analogous spliceosome complex mediates the excision of a rare subset of evolutionary conserved introns that exhibit non-canonical consensus sequence, referred to as minor-class introns (Hall and Padgett 1994). The minor-class intron spliceosomes are low abundant and formed by the distinct but functionally analogous snRNPs U11, U12, U4atac, and U6atac together with the U5 snRNPs, which is a particle shared by both machineries. The much less frequent minor-class introns coexist with neighboring canonical major-class introns in a gene. The two spliceosome machineries undergo comparable dynamic rearrangements, with the main differences occurring at the early stages of intron recognition rather than during catalysis (Patel and Steitz 2003). The minor-class splicing follows the same two-step reactions and formation of a lariat intermediate as the major splicing. U11 base-pair with the characteristic longer and constrained consensus sequence at the 5′-ss of the minor class introns, whereas U12 base-pair with the BPS. The secondary structure of U11 and U12 mimics that of U1 and U2 snRNAs, respectively. Minor-class introns lack the Py-tract. Analogous to the major pathway, the U4atac snRNP chaperone U6atac into the spliceosome, preventing U6atac interaction with U12 and the 5′-ss before their helicase-dependent unwinding. Upon unwinding, U4atac is released, followed by rearrangements that permit catalytic activation (Patel and Steitz 2003).

Initial recognition and pairing of the 5′ and 3′ splice sites depend very much on the size of the intron and the distance between the splice sites affects the efficiency in which spliceosome assembles (Fox-Walsh et al. 2005). Because splice sites are recognized across an optimal nucleotide length, depending on the size of the intron or the flanking exons, the splice sites are recognized across the intronic or exonic segments, known as exon and intron definition models (Berget 1995; De Conti et al. 2013). When exons are small and introns are long, the splicing machinery forms across exons, whereas in genome architecture, where exons are large and introns are small, such as observed in lower eukaryotes, the intron definition prevails. In the human genome, where the majority of exons are short and introns are long, it is likely that the vast majority of splice sites are recognized across the exon (Fig. 2.2). The same splicing complexes formed across exon operate on intron definition in terms of composition and structure, and both exon and intron definition models may co-occur within the same pre-mRNA (De Conti et al. 2013; Li et al. 2019).

2.2 Alternative Splicing

Whereas some exons are constitutively spliced, that is, they are present in every mRNA produced from a given pre-mRNA, others are alternatively spliced to generate variable forms of mRNA from a single pre-mRNA. By shifting the exon usage, alternative splicing (AS) notably enhances the transcriptome and cellular proteome (Pan et al. 2008). Alternatively, spliced gene products may have related, distinct, or even opposing functions as well as non-functional properties, which may lead to the regulation of its expression through nonsense-mediated mRNA decay (NMD) or nuclear sequestration and turnover. Consequently, AS represents an important level of

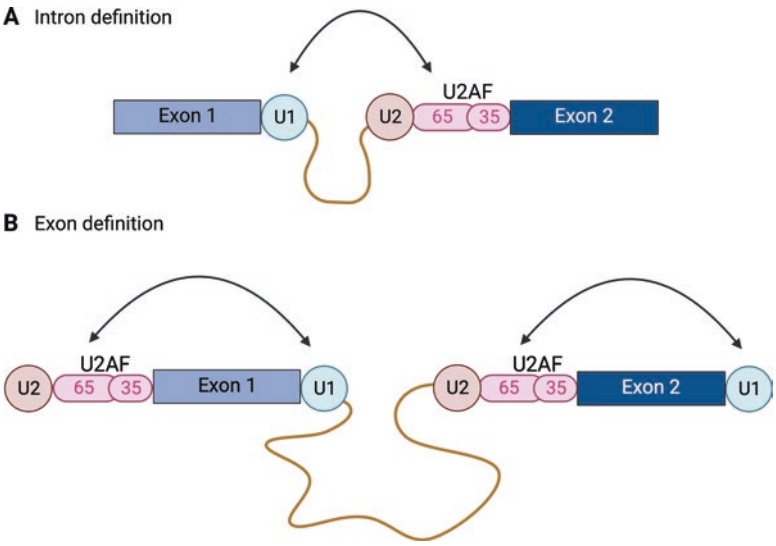

Fig. 2.2 *Intron and exon definition models*. (**a**) Intron definition model: the U1 and U2 snRNPs and U2AF interact to allow the pairing of the splice sites across an intron when the intron is short (<250 bp). (**b**) Exon definition model: the pairing of the splice sites occurs across the exon when separated by a long intron (>250 bp). (Made in ©BioRender – biorender.com)

regulation in gene expression and plays a critical role in biological processes such as development, cell differentiation, and response to environmental cues. AS frequency increases with species complexity and it has been proposed to be a driver of phenotypic complexity evolution in mammals. Accordingly, significative higher frequencies of AS events are observed in brain tissues throughout vertebrate species where regulation of these splicing events has been associated with evolutionary changes contributing to nervous system development (Barbosa-Morais et al. 2012; Merkin et al. 2012).

Virtually all human multi-exon genes undergo alternative splicing (Pan et al. 2008; Wang et al. 2008). The main types of AS events are exon skipping, mutually exclusive exon, alternative 5′ and 3′ splice site selection, and alternatively retained introns (Nilsen and Graveley 2010). Exon skipping takes place when a particular cassette exon is spliced out of the mature message. In a mutually exclusive exon AS event, one out of two exonic regions of a pre-mRNA is included in the final transcript when the other is excluded, and vice versa, so that these exons never coexist in the same product. Alternative splice site selection occurs when spliceosome recognizes and pairs with cryptic splice sites, resulting in the alternative splice site cleavage on the nearby exon. The fifth type of AS event, the retained intron, is the process by which a particular intron remains unspliced in the final transcript, and as a result, it triggers the downregulation of the transcript through NMD or nuclear sequestration and turnover (Braunschweig et al. 2014). In addition to the common types of AS, alternative promoter usage and alternative cleavage and polyadenylation yield different types of alternative transcript events, such as alternative first exon (AFE), tandem 3′ untranslated region (UTR), and alternative last exon (ALE).

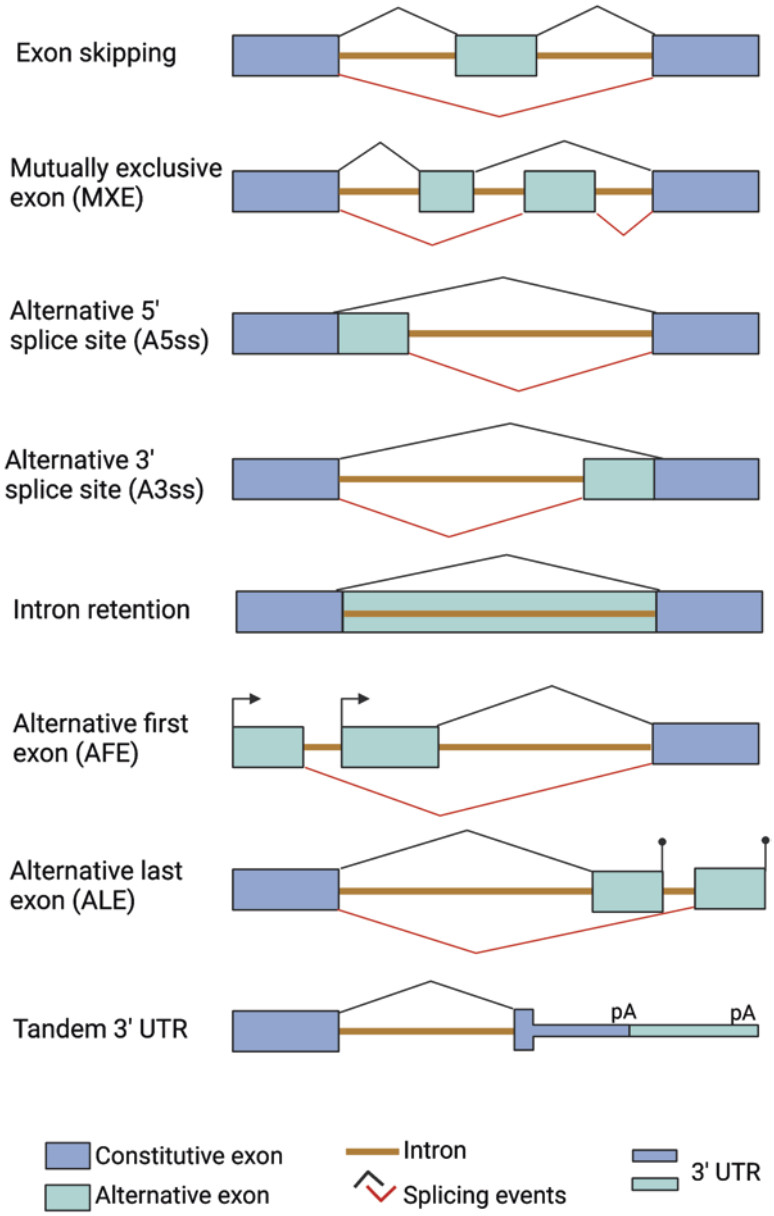

Fig. 2.3 *Types of alternative splicing (AS) events*. (Made in ©BioRender – biorender.com)

In the first case, alternative promoter usage gives rise to mRNA isoforms with distinct 5′ UTR, and in the second case, the usage of alternative polyadenylation sites gives rise to transcripts with shorter or longer 3′ UTR and with distinct terminal exons (Wang et al. 2008) (Fig. 2.3). A combination of different modes of AS is often observed throughout a single precursor mRNA of multiexon genes.

In humans, the splice site sequences are highly degenerated (Sheth et al. 2006) and often not sufficient to define exon–intron boundaries. In addition, sequences that match the short consensus splice site signals are commonly found throughout the introns. In order to help the splice site selection, exons and their nearby intronic regions contain a variety of additional splicing regulatory elements (SREs). If they enhance exon inclusion, these elements are called exon splicing enhancers (ESE) or intronic splicing enhancers (ISE), depending on whether they are present in exonic or intronic regions. If they tend to repress exon inclusion, these elements are called exon or intron splicing silencers (ESS or ISS, respectively) (Zhang et al. 2008).

These cis-acting sequences within the pre-mRNA influence splicing through the binding of specific RNA-binding non-spliceosomal regulatory proteins, which either promote or hinder the spliceosome activity on the adjacent splicing sites (Cartegni et al. 2002).

The requirement for additional cis-acting and trans-acting elements to stabilize and target specific sites introduces another layer of complexity in the regulation of the splicing machinery and provides an important window for variations and diversity.

Two major classes of widely expressed trans-acting factors, namely the SR proteins and the heterogeneous ribonucleoproteins (hnRNPs), are involved in recognizing and binding the cis-elements within the RNA (Long and Caceres 2009; Martinez-Contreras et al. 2007).

Proteins of the SR family contain one or two N-terminal RNA recognition motif (RRM), which mediate binding to RNA and a C-terminal arginine-serine-rich (RS) domain, involved mainly in protein–protein interaction. They play an important role as general splicing factors and as regulators of alternative splicing (Manley and Krainer 2010). The hnRNPs form a group of structurally diverse RNA binding proteins involved in different stages of the RNA metabolism (Geuens et al. 2016).

Several reports describe the antagonistic function of SR and hnRNP proteins on alternative splicing (Cáceres et al. 1994). Typically, splicing enhancers are recognized by a member of the SR family, whereas hnRNP recognizes splicing silencers (Cartegni et al. 2002). SR sequence motifs are enriched in exonic sequence (Liu et al. 1998). When SR proteins are bound to ESEs they favor exon inclusion and prevent exon skipping. They can also promote exon definition by directly recruiting the splicing machinery through their RS domain and antagonizing nearby silencer elements.

On the other hand, hnRNP sequence motifs are enriched in introns. hnRNP represses splicing by directly antagonizing the recognition of splice sites or interfering with the binding of proteins bound to enhancers. Various hnRNPs regulate alternative splicing by stimulating exon skipping or intron retention (Fig. 2.4). To add an additional layer of complexity to this network, some SR protein can be implicated in splicing silencing when associated with introns, whereas some hnRNP can inhibit splicing from exonic locations. Specifically, SR and hnRNP protein activities differ depending on their position relative to the regulated splice, in which some SR protein can repress splicing, whereas hnRNP can enhance splicing depending on their position relative to the regulated site (Erkelenz et al. 2013; Matera and Wang 2014).

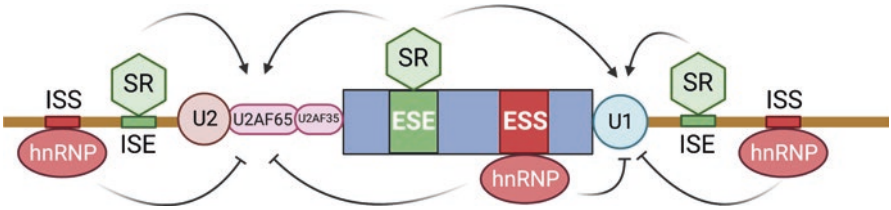

Fig. 2.4 *The antagonistic role of SR proteins and hnRNPs in splicing*. In general, when bound to enhancer sequences, SR proteins positively influence splicing promoting exon inclusion, exon definition, and recruiting the splicing machinery. hnRNPs, conversely, bind to silencing elements, inhibiting the recognition of splice sites, and interfering with the binding of proteins to enhancers. ESE: exonic splicing enhancer; ESS: exonic splicing silencer; ISE: intronic splicing enhancer; ISS: intronic splicing silencer. (Made in ©BioRender – biorender.com)

In the end, the decision of whether a specific site is selected or if a particular exon or intron is included or excluded from the final transcript is defined by the combinatorial interplay of positive and negative regulatory signals present in the RNA, the ultimate complexes formed by the trans-acting factors assembled on these regulatory sequences and how they influence the splicing machinery on the nearby splice sites. On top of that, variations in the relative concentrations of the antagonistic trans-acting elements may affect splice-site choice by tipping the balance in favor of different outcomes. Thus, the ratios of these antagonistic factors are likely to define a cellular code for establishing cell-specific patterns of splicing in multiple genes (Smith and Valcárcel 2000).

2.3 Mechanisms of Alternative Splicing Regulation

A much higher level of complexity is added to alternative splicing regulation when considering the fact that most splicing evens occur co-transcriptionally. The coupling of splicing and transcription implies a tight integration of alternative splicing with other gene regulatory pathways (Braunschweig et al. 2013).

2.3.1 RNA Architecture and Secondary Structures

RNA architecture impacts splicing outcomes. Secondary structures on the pre-mRNA can influence the accessibility of splice sites or cis-acting elements. These secondary structures are formed by intramolecular base-pairing and impact splice site selection positively or negatively. Thus, RNA folding may be regarded as an important component of AS regulation (McManus and Graveley 2011; Warf and Berglund 2010).

The mechanisms by which these structures influence splicing may involve local- and long-range interactions within the RNA molecules. Local pairing forms RNA structures that may prevent the binding of appropriate regulatory factors to the single-stranded molecule, which can happen when secondary structures are formed onto cis-regulatory elements or in its vicinity. Similarly, when secondary structures overlap splicing signals, such as 5′-ss, 3′-ss, BPS, and Py-tract, it can hamper recognition and assembly of spliceosome complexes onto these sites.

Conversely, long-range interactions place distant sequences into proximity, which promotes the looping out of specific regions of the pre-mRNA. These regions may contain cassette exons, a stretch of exonic and intronic sequences, or functionally active cis-acting elements that are drawn to alternative splicing regulation. Besides, approximation of distant regulatory elements and cognate factors to target exons may favor alternative splice site selection. The role of alternative competing RNA folding in the choice of alternative splice sites and mutually exclusive exons was first described for the drosophila *Dscam* gene, but similar modes of regulation have been proposed for human genes (Pervouchine et al. 2012).

While RNA secondary structure impacts splicing, its formation is condition-dependent and may also be subject to regulation. For instance, RNA helicases can unwind these structures and consequently regulate pre-mRNA splicing. Also, transcription rate can influence the folding of the RNA during synthesis and ultimately if slow-folding structures have enough time or not to assemble on the nascent pre-mRNA before splicing occurs.

2.3.2 Coupling Transcription to Alternative Splicing

Besides the interaction of snRNPs and splicing regulatory factors with pre-mRNA, the high-fidelity process of splice site selection also requires transcriptional machinery as well as chromatin modifiers and remodelers (Luco et al. 2011). Each step of transcription, namely initiation, elongation, and termination, contributes to how the nascent pre-mRNA is processed.

The impact of promoter usage in alternative splicing was first demonstrated on experimental models in which an artificial minigene under the regulation of different promoters exhibited a different pattern of alternative transcript. This led to the conclusion that promoters may contribute to AS by recruiting different molecules to the transcriptional complex, which in turn participate in the splicing regulation. The primary implication of these findings was that cell-specific AS may not simply result from the differential abundance of SR proteins but also from a more complex process involving cell-specific promoter occupation. Unlike the minigene experimental model, most of the genes are regulated by a single promoter in nature. In this case, the differential occupancy of the promoters by a variety of transcription factors and co-activators impacts AS. The promoter itself is responsible for recruiting splicing regulator factors to the site of transcription, possibly through the interaction with transcription factors bound to the promoter or transcriptional enhancers. Also,

some of the effects of promoters on pre-mRNA splicing are mediated by proteins that function as dual transcription and splicing factors (Kornblihtt 2005).

The RNA polymerase II (RNAPII) has a major impact on alternative splicing, and its largest subunit C-terminal domain (CTD) plays a central role in coupling the two processes. In mammals, the CTD domain comprises 52 heptad repeats (YS_2PTS_5PS), subject to extensive phosphorylation. Phosphorylated CTD serves as a binding module for multiple mRNA processing factors, including splicing factors. Phosphorylation of the CTD repeats on serine 5 residues are essential for capping enzyme recruitment, whereas phosphorylation on serine 2 facilitates recruitment of cleavage and polyadenylation factors at 3′ ends of the RNA. Particularly, serine 2 phosphorylation was shown to be essential for the integration of transcription and splicing. In addition, the mediator complex, known to facilitate the interaction between the transcription pre-initiation complex on promoters with distant transcription regulatory factors bound to enhancers, also contacts splicing factors (David and Manley 2011).

The regulation of the RNAPII elongation rate constitutes another mechanism in which transcription affects AS. The RNAPII elongation rate is governed by the rates of RNA synthesis and translocation of the enzyme interspersed by acceleration, deceleration, backtracking, pausing and release, and sometimes with premature termination, which may occur while the RNAPII transcription elongation complex travels along a given gene. The use of RNAPII rate mutants to investigate the impact of elongation rate on genome-wide alternative splicing in human cells demonstrated that both slow and fast transcription changed the alternative splicing of thousands of exons. Slower transcription rates mainly contribute to cassette exons' inclusion, whereas a faster transcription leads to their skipping from mature mRNAs (Saldi et al. 2016).

Two models were proposed to illustrate the coupling of transcription and splicing processes, namely the recruitment and the kinetic models. In the recruitment model, a change in promoter architecture results in the recruitment of splicing factors to the transcription machinery that in turn impact the splicing of the nascent RNA. In the kinetic model, the change in promoter architecture affects the elongation rate of the RNAPII, such that there is more or less time for splice sites or other splicing signals flanking the alternative exons to be recognized by trans-acting factors. Thus, this model predicts that elongation rate modulates competition between splice sites and cis-regulatory elements. Such as, if there is a cassette exon flanked by weak upstream 3′-ss and a strong downstream 3′-ss, a lower transcription rate will favor the usage of the upstream site and, consequently, the inclusion of the cassette exon. On the other hand, acceleration of transcription will favor the usage of the downstream site, resulting in exon skipping (Fig. 2.5a). In addition, a lower transcription rate can also favor exon skipping if, for example, an ISE is displayed upstream of a cassette exon. And as mentioned before, the elongation rate can also influence RNA folding events, which contributes to the AS outcome (Braunschweig et al. 2013).

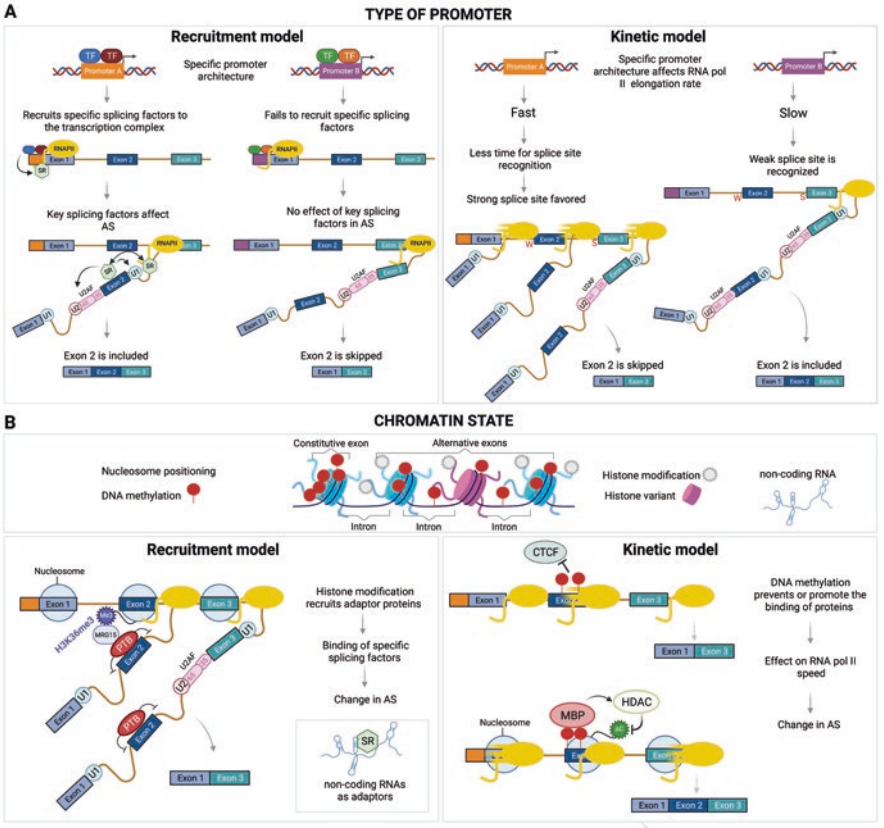

Fig. 2.5 *Promoter and chromatin features affect AS through recruitment and kinetic models*. (**a**) Recruitment model (left): the type of promoter will determine the transcription factors (TFs) that will be recruited to the transcription initiation complex. Promoter A recruits TFs that interact with splicing factors (for example, SR proteins) that will enhance splicing of the alternative exon. The TFs on Promoter B do not interact with splicing factors, affecting the splicing outcome differently from Promoter A. Kinetic model (right): the type of promoter will determine RNAPII elongation rate. Promoter A determines a high elongation rate of the RNAPII, allowing less time to the CTD-associated splicing factors to recognize weak (W) splice sites. Higher RNAPII elongation rate favors recognition of strong (S) splice site. Promoter B determines a slower RNAPII elongation rate, allowing for weak splice site recognition and exon inclusion. (**b**) The chromatin architecture is dictated by many epigenetic layers, such as nucleosome positioning, DNA methylation, histone modification, histone variants, and also non-coding RNAs. All these aspects of chromatin influence alternative splicing. Recruitment model (left): histone modification (such as H3K36me3) recruits an adaptor protein (MRG15) which in turn recruits a splicing factor (PTB, an hnRNP protein with repressor activity) that will affect splicing of the alternative exon. Non-coding RNAs, such as lncRNAs, can act as a scaffold to recruit or sequestrate specific splicing factors. Kinetic model (right): chromatin features affecting the RNAPII elongation rate. For example, DNA methylation prevents CTCF binding, which acts as a roadblock for RNAPII. Without the binding of CTCF, elongation is accelerated and disfavors alternative exon inclusion. DNA methylation on alternative exons can also recruit methyl-binding proteins (MBP) such as MeCP2. MBP recruits histone deacetylase complex (HDAC), leading to less permissive chromatin and slower RNAPII elongation, favoring alternative exon inclusion. (Made in ©BioRender – biorender.com)

2.3.3 *Epigenetic Control of Alternative Splicing*

Transcription rate is influenced by the binding of transcription factors and co-factors to promoters and regulatory sequences on the DNA, which in turn is shaped by the chromatin structure and the intricate interplay of epigenetic modifications.

The various layers of epigenetic control – DNA methylation, nucleosome positioning, histone modifications, histone variants, chromatin remodeling factors, and non-coding RNAs – are involved in the regulation of AS.

The 5′ cytosine methylation, deposited by DNA methyltransferases (DNMTs) on CpG dinucleotides, corresponds to an essential epigenetic modification on the DNA that influences gene expression patterns across the genome. While deposition of the 5-methylcytosine (5mC) at promoter regions exerts an inhibitory effect on gene expression, its presence on the genes bodies positively affects transcription of the related genes, besides preventing spurious transcription initiation from cryptic internal promoters. 5mC are enriched at exons and especially at splice sites when compared to flanking introns. Moreover, DNA methylation is less abundant in alternatively spliced exons than in constitutive exons (Lev Maor et al. 2015). DNA methylation can either enhance or silence exon recognition. There are three different mechanisms by which DNA methylation regulates AS. DNA methylation can prevent the DNA binding protein, CTCF, from interacting to its binding site, counteracting its function as a roadblock for RNA pol II, culminating in increased elongation rate and exon skipping. Another mechanism involves the binding of methyl-binding proteins, such as MePC2. Binding of MePC2 to methylated DNA triggers recruitment of histone deacetylase complex (HDAC), local hypoacetylation, and consequent RNA pol II pause, favoring exon inclusion (Lev Maor et al. 2015). In a third mechanism, DNA methylation on alternative exons induces the H3K9me3 histone modification, which in turn anchors Heterochromatin Protein 1 (HP1) isoforms HP1α and HP1β. HP1α and HP1β act as adaptor proteins for the recruitment of splicing factors. The presence or absence of HP1 and its associated splicing factors determine whether a cassette exon is included or excluded from the transcript (Yearim et al. 2015).

Nucleosomes are the basic units of chromatin, comprised of an octamer of histones. The DNA wrapped around the nucleosome is approximately 147 nt in length, which is also the average size of exons in mammals. Nucleosomes are enriched in exons, and because of that, they are determinants in exon definition. Moreover, exons that lay between long introns present a higher nucleosome positioning than exons separated by small introns. These facts suggest that nucleosomes may act in both protecting and defining exons (Luco et al. 2011). Alternatively, included exons flanked by weak splice sites present higher nucleosome density than excluded exons, pointing to nucleosome function in AS. Nucleosomes influence AS by acting as a barrier that controls RNA pol II density at exons, lowering transcription rate and favoring the inclusion of cassette exons (Saldi et al. 2016).

Post-translational modification of histones is an important determinant of chromatin. Histone modification such as H3 lysine 36 trimethylation (H3K36me3) and H3 lysine 9 trimethylation (H3K9me3) also influence alternative splicing (Luco et al. 2011). H3K36me3 is highly abundant in actively transcribed genes and is deposited by the methyltransferase SETD2. Dysfunction in this enzyme leads to changes in alternative splicing events. H3K36me3 mark influences splicing by anchoring the chromatin-binding protein, MRG15, which recruits the polypyrimidine binding protein (PTB), a splicing repressor. Therefore, the levels of H3K36me3 will determine the ultimate effect of PTB, favoring exon inclusion or skipping (Luco et al. 2010). Likewise, the H3K9me3 mark also promotes exon skipping, in this case, by recruiting the Heterochromatin protein 1 (HP1). HP1 isoforms HP1α and HP1β act as adaptors to recruit the protein SRSF3. The SRSF3 protein is a member of the SR family that functions as a splicing silencer hindering the inclusion of cassette exons.

As previously mentioned, the H3K9me3 mark recruits HP1, but different from HP1α and HP1β, the isoform HP1γ has a different mode of action. HP1γ binds to the H3K9me3 marks in the gene's coding region and simultaneously associates with the pre-mRNA. Interaction of HP1γ with the nascent pre-mRNA slows RNAPII and consequently the elongation rate. Decrease in the RNAPII elongation rate allows time to recruit splicing factors and cassette exon inclusion (Saint-André et al. 2011).

Additionally, histone variants play a role in AS. The variants H3.3A, H3.3B, H2a.V, and the H3-histone chaperone Asf1 play a role in the processing of histone RNAs itself. It has been described in human lung fibroblasts that deposition of the linker histone variant H1.5 at the splicing sites of short exons constitutes a mark responsible for considerable stalling of RNAPII. Again, decreased RNAPII elongation rate facilitates the inclusion of alternative exons (Glaich et al. 2019). Similarly, BRM – the ATPase subunit of the switch/sucrose non-fermenting, SWI/SNF – a chromatin remodeling factor, facilitates the inclusion of alternative exons by interacting with RNAPII and inducing its pause (Batsché et al. 2006).

In conclusion, the dynamics of the chromatin imposed by DNA methylation, nucleosome positioning, and histone modifications control alternative splicing through the mechanisms portrayed by the recruitment and the kinetic models described earlier (Braunschweig et al. 2013) (Fig. 2.5b).

Ultimately, non-coding RNAs are epigenetic regulators that can affect the chromatin structure and also alternative splicing. Long non-coding RNAs (lncRNAs), such as Metastasis-associated lung adenocarcinoma transcript 1 (MALAT1), taurine upregulated gene 1 (TUG1), and Gomafu, act as a scaffold for the binding of splicing factors, affecting their function in splicing (Ramanouskaya and Grinev 2017). For example, Malat1 associates with SR proteins such as SRSF1, SRSF2, and SRSF3, and its deletion changes the alternative splicing pattern of genes related to tumorigenesis (Zhang et al. 2020).

2.3.4 Coupling Cell Signaling to Alternative Splicing

Alternative splicing also responds to the constant changes in the cell's physiological or pathological conditions (Kornblihtt et al. 2013). In a seminal work, König and colleagues described the mechanism by which splicing can be coupled with signal transduction. The authors found signal-responsive elements within the exonic v5 sequence of the *CD44* pre-mRNA. These ESS and ESE elements modulate alternative splicing in a cell type-specific and inducible manner in response to Ras signaling pathway (König et al. 1998). The mechanism involved Ras-Raf-MEK-ERK activation and SAM68 phosphorylation by ERK. Phosphorylated SAM68 bound to the ESS element resulting in the inclusion of the v5 exon through mechanisms that include impairing the splicing repressor hnRNP A1 from binding to the ESS element, recruiting proteins that promote spliceosome assembly, interacting with chromatin remodeling complex, and decreasing RNAPII elongation rate, and thus favoring the use of weak splice sites (Frisone et al. 2015; Lynch 2007). Together with the protein interactions with the ESE element that enhance splicing, exon v5 is included in the mature mRNA in response to an extracellular signal (Shin and Manley 2004). The TGF-β signaling also controls alternative splicing of the *CD44* pre-mRNA, leading to the expression of the cancer aggressiveness-related isoform CD44v6 (Tripathi et al. 2016).

There are other examples of signaling pathways controlling alternative splicing. The mechanism applied in the majority involves regulation of the SR protein activity through phosphorylation and dephosphorylation events. Such as, Fas receptor activation includes the activity of the phosphatase PP1, which dephosphorylates SR proteins. Altered phosphorylation of the SR splicing factors results in a switch of BCL-X and Caspase 9 transcripts from anti- to pro-apoptotic isoforms. Likewise, phosphoinositide signaling leads to dephosphorylation of the SR factors SRSF10. Dephosphorylated SRSF10 interacts with U1 snRNPs interfering with its 5′-ss recognition and impairing splicing. Also, AKT pathway mediates alternative splicing in response to epidermal growth factor (EGF) signaling through phosphorylation of SR proteins (Kornblihtt et al. 2013).

2.4 Missplicing and Disease

Aberrant splicing causes diseases. According to the Human Gene Mutation Database (HGMD), mutations affecting splicing account for one-third of all disease-causing mutations. The mutations trigger aberrant splicing by mechanisms that involve either disruption of splicing signals or cis-acting regulatory elements on the RNA or interference with the function of the trans-acting factors that act on the RNA.

Mutations in regulatory sequences that affect alternative splicing are a widespread cause of human hereditary disease and cancer. These cis-acting mutations disrupt the splicing code in different ways. It can affect splice sites (5′- and 3′-ss),

Py-tract, or BPS or create cryptic splicing signals. It can alter sequences that overlap with the secondary structure of the RNA, hampering its formation or creating folding that is not usually there. Moreover, cis-acting mutations can result in the loss or gain of function of splicing enhancers (ESE, ISE) or silencers (ESS, ISS). The consequence of these alterations is the aberrant splicing of the involved genes due to exon skipping, intron retention, activation of cryptic sites, and the altered ratio of skipping/inclusion of cassette exons.

The following selected examples illustrate disease-associated splicing alterations caused by the different types of *cis*-acting mutations. (i) Familial dysautonomia (FD) is a recessive genetic disorder characterized by a point mutation in the vicinity of the 5′-ss on intron 20 of the *IKBKAP* gene. The altered splice site impairs the recognition and base-pairing of the U1 snRNP at the 5′-ss resulting in exon skipping. (ii) Beta-thalassemia is an inherited disorder characterized by reduced expression of the hemoglobin beta chain and severe anemia. One form results from a point mutation present in the intron 1, which generates an alternative 3′-ss in the *HBB* gene. This cryptic site is preferentially used and results in an aberrant isoform. (iii) The frontotemporal dementia and Parkinsonism linked to Chr.17 (FTDP-17) is a neurodegenerative disorder that can be caused by a point mutation in the exon 10 of the *MAPT* gene. This mutation affects an ESS leading to increased inclusion of the exon. (iv) Familial partial lipodystrophy type 2 (FPLD2) is a rare metabolic condition characterized by point mutations in 5′-ss of the *LMNA* gene, resulting in intron retention and consequent regulation of its transcripts expression through NMD (Daguenet et al. 2015; Scotti and Swanson 2016).

Similar to cis-acting mutations, genetic variations that naturally occur in the population, such as single nucleotide polymorphisms (SNPs), also affect the efficiency of alternative splicing. An example of an allele-dependent expression of alternative isoform is observed for the major histocompatibility complex, class II, DQb 1 (HLA-DQB1) gene. The ratio of *DQB1* exon 4 inclusion in the final transcript is determined by differential recognition of the upstream 3′-ss during the early steps of spliceosome assembly. The differential recognition of the 3′-ss results from differences in the RNA sequence due to the SNPs mapped to this region, directly affecting the BPS and Py-tract (Královičová et al. 2004).

In addition to mutations affecting splicing signals on the pre-mRNA, mutations in genes coding for *trans*-acting factors also cause disease. Unlike the *cis*-acting mutations that only affect the compromised gene, trans-acting mutations alter the function of proteins implicated in the splicing machinery and thus convey a pleiotropic effect on large sets of genes.

Disease-associated trans-acting mutations affect genes involved in UsnRNP biogenesis and assembly or formation of UsnRNP aggregates, spliceosome assembly (core spliceosome mutations), and splicing regulation (SR, hnRNP, and RNA binding proteins).

Spinal muscular atrophy (SMA), Clericuzio-type poikiloderma with neutropenia (PN), Retinitis pigmentosa (RP), and Alzheimer's disease are examples of disorders associated with mutations that interfere with the function of proteins involved in UsnRNP biogenesis. Prader-Willi syndrome and RP are examples of disorders

associated with mutations that interfere with core spliceosome- and splicing factor protein. At the same time, amyotrophic lateral sclerosis (ALS), autism disorder, and Huntington's disease have been associated with deregulated expression of splicing factors and RNA-binding proteins.

The following selected examples illustrate disease-associated trans-acting mutations involved in each of the mechanisms described. (i) Clericuzio-type poikiloderma with neutropenia (PN) rare autosomal recessive disease associated with mutations in the *C16orf57* gene. This gene encodes a protein involved in the correct processing of the U6 snRNA. Cells from patients carrying these mutations have higher levels of U6 snRNA degradation. (ii) RP are inherited degenerative disorders of the retina associated with mutations in multiple genes, among them the pre-mRNA processing factors *PRPF31*, *PRPF8*, *PRPF6*, *PRPF3*, and the RNA helicase *SNRNP200/BRR2*. These genes code for components of the spliceosome complex involved in complex rearrangement and catalysis. (iii) ALS is a common motor/neurodegenerative disease caused by mutations in various genes, such as the ones coding for the RNA-binding protein FUS and TDP-43. Their altered function affects snRNA abundance and, consequently, pre-mRNA splicing (Daguenet et al. 2015).

2.5 Splicing and Cancer

In the context of cancer, mutations in the cis-acting regulatory sequences on the RNA often generate aberrant tumor-associated isoforms that contribute to some aspects of tumorigenesis. Additionally, cis-acting mutations can contribute to activate oncogenic isoforms or inactivate tumor suppressor transcripts. On a broader scale, somatic mutations in genes encoding components of the splicing machinery are also frequently observed and are related to global splicing abnormalities of cancer transcriptomes (Dvinge et al. 2016).

Recurrent somatic mutations in core spliceosome and splicing factor coding genes were first discovered in hematological malignancies like myelodysplastic syndromes (MDS), acute myeloid leukemia (AML), and chronic lymphoblastic leukemia (CLL) (Graubert et al. 2012; Papaemmanuil et al. 2011; Quesada et al. 2012; Wang et al. 2011; Yoshida et al. 2011), and later identified with high frequency in a variety of solid tumors, such as uveal melanoma (Harbour et al. 2013), lung adenocarcinoma (Imielinski et al. 2012), breast (Maguire et al. 2015; Stephens et al. 2012), and pancreatic cancer (Biankin et al. 2012). These findings were the first direct genetic link between dysfunction of splicing machinery and cancer, and defects in this machinery have been proposed as leukemogenic pathways (Maciejewski and Padgett 2012).

Interestingly, these mutations are heterozygous and mutually exclusive, indicating that cells may tolerate only partial deviation from normal splicing. In fact, cells carrying splicing factor mutations are sensitive to genetic or pharmacological perturbation of splicing (Fei et al. 2016; Obeng et al. 2016; Seiler et al. 2018; Shirai et al. 2017; Zhou et al. 2015).

The most frequently reported splicing mutations in cancer occur in four genes, namely *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2*. At least for *SF3B1*, *U2AF1*, and *SRSF2*, mutations affect highly restricted residues within the proteins, suggesting again (or alteration) of function phenotype, whereas *ZRSR2* mutations are widespread throughout the protein and follow a loss of function pattern. These mutations affect splicing by interfering with 3′-ss recognition mediated by *SF3B1; U2AF1* on U2-type introns or by *ZRSR2* on U12-type introns as well as with exon recognition mediated by SRSF2 (Dvinge et al. 2016).

Among the less frequently mutated splicing factors genes are *SF1*, *U2AF2*, *SF3A1*, *PRPF40B* (Yoshida et al. 2011), *PRPF8*, *LUC7L2*, *HCF1*, *SAP130*, *SRSF6*, *SON*, and *U2AF26* (Makishima et al. 2012). Together with *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2*, these genes encode multiple components or associated factors from the spliceosome complexes E/A (Fig. 2.6). Although less frequently mutated, these

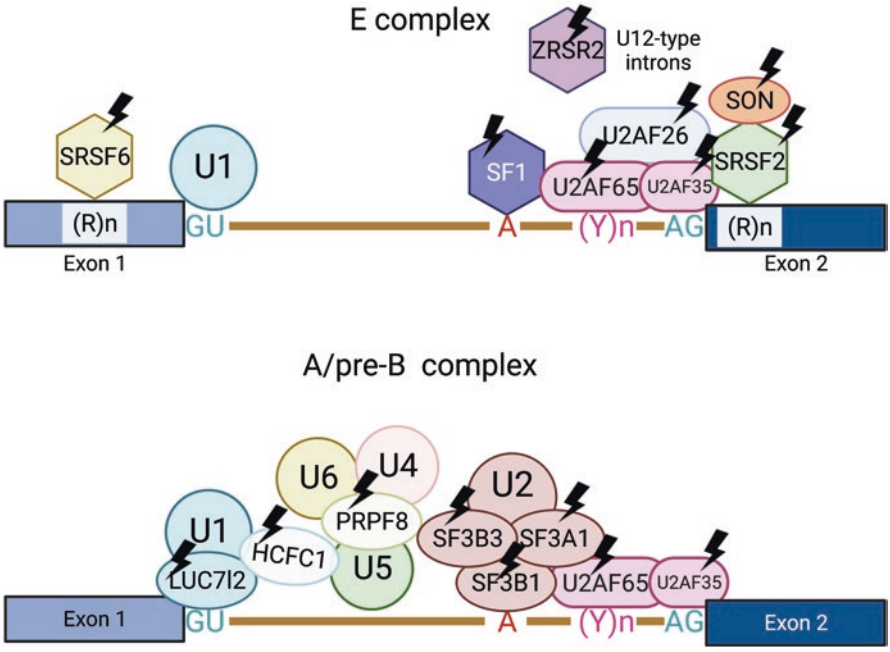

Fig. 2.6 *Core spliceosome and splicing factor proteins affected by somatic mutation in cancer*. A lightning symbol represents the affected components. E complex: U2AF35 binds to AG dinucleotide on the 3′-ss of the intron, while U2AF65 binds to the Py-tract, represented by (Y)n. SF1 binds to the BPS, represented by an A. U2AF26 interacts with U2AF to perform essential functions in splicing. ZRSR2 acts on 3′-ss recognition of U12-type introns. Arginine/serine-rich splicing factors SRSF2 and SRSF6 bind to polypurine sequences ((R)n) in the exon. SRSF2 interacts with U2AF65. SON, a recently discovered spliceosomal gene, interacts with SRSF2 and mediates constitutive splicing of weak splice sites. A/pre-B complex: SF1 is replaced by U2 snRNP along with its components SF3A1, SF3B1, and SF3B3. LUC7L2 is associated with the U1 snRNP on 5′-ss. PRPF8 plays an essential role in the interaction among U4/U6/U5 snRNPs, while HCFC1 contributes to the U1/U5 interaction. (Made in ©BioRender – biorender.com)

genes participate in the same molecular pathway as the frequently mutated genes, indicating that the impairment of the pathway rather than the individual molecule is important for carcinogenesis.

In addition to mutations affecting *trans*-acting factors, the differential expression of splicing regulatory factors and altered post-translational modification of these proteins are strongly associated with splicing abnormalities and transformation. Mounting evidence suggests that these factors can act as both oncoproteins and tumor suppressors. Among the growing list are the SR proteins, hnRNPs, and other splicing factors such as SRSF1, SRSF3, SRSF6, SRSF10, hnRNP A1, hnRNP A2, hnRNP A2/B1, hnRNP H, hnRNP K, hnRNP A2 hnRNP M, PRPF6, PTB QKI, RBFOX2, RBM4, RBM5, RBM6, and RBM10 (for review, see Dvinge et al. 2016; Grosso et al. 2008).

The first mechanistic evidence that deregulated splicing factor expression resulted in the malignant transformation was demonstrated for the SR factor, SRSF1. SRSF1 is upregulated in several tumors, and this is sufficient to affect the alternative splicing of the BIN1 tumor suppressor and the MNK2 and S6K1 kinases. The resulting isoform of BIN1 has no tumor suppressor activity, whereas those of MNK2 and S6K1 have shown oncogenic properties (Karni et al. 2007).

There is an overlap between the splicing factors with altered expression in cancer, which are also mutated in hematological malignancies, such as *U2AF1*, *SRSF2*, *SFRS6*, and *SF1* (Grosso et al. 2008), indicating that disturbing the function of these proteins at any level might contribute to disease.

A major challenge in research today is to associate the mutations and aberrant expression/activity of the splicing factors with specific downstream splicing changes.

References

Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, Slobodeniuc V, Kutter C, Watt S, Çolak R, Kim T, Misquitta-Ali CM, Wilson M, Kim P, Odom DT, Frey B, Blencow B (2012) The evolutionary landscape of alternative splicing in vertebrate species. Science 338(80):1587–1593

Batsché E, Yaniv M, Muchardt C (2006) The human SWI/SNF subunit Brm is a regulator of alternative splicing. Nat Struct Mol Biol 13:22–29

Berget SM (1995) Exon recognition in vertebrate splicing. J Biol Chem 270:2411–2414

Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch AM, Wu J et al (2012) Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 491:399–405

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ (2013) Dynamic integration of splicing within gene regulatory pathways. Cell 152:1252–1269

Braunschweig U, Barbosa-Morais NL, Pan Q, Nachman EN, Alipanahi B, Gonatopoulos-Pournatzis T, Frey B, Irimia M, Blencowe BJ (2014) Widespread intron retention in mammals functionally tunes transcriptomes. Genome Res 24:1774–1786

Cáceres JF, Stamm S, Helfman DM, Krainer AR (1994) Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. Science 265(80):1706–1709

Cartegni L, Chew SL, Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 3:285–298
Cordin O, Beggs JD (2013) RNA helicases in splicing. RNA Biol 10:83–95
Daguenet E, Dujardin G, Valcárcel J (2015) The pathogenicity of splicing defects: mechanistic insights into pre- mRNA processing inform novel therapeutic approaches. EMBO Rep 16:1640–1655
David CJ, Manley JL (2011) The RNA polymerase II C-terminal domain-a new role in spliceosome assembly. Transcription 2:227–231
De Conti L, Baralle M, Buratti E (2013) Exon and intron definition in pre-mRNA splicing. Wiley Interdiscip Rev RNA 4:49–60
Dvinge H, Kim E, Abdel-Wahab O, Bradley RK (2016) RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer 16:413–430
Erkelenz S, Mueller WF, Evans MS, Busch A, Schöneweis K, Hertel KJ, Schaal H (2013) Position-dependent splicing activation and repression by SR and hnRNP proteins rely on common mechanisms. RNA 19:96–102
Fei DL, Motowski H, Chatrikhi R, Prasad S, Yu J, Gao S, Kielkopf CL, Bradley RK, Varmus H (2016) Wild-type U2AF1 antagonizes the splicing program characteristic of U2AF1-mutant tumors and is required for cell survival. PLoS Genet 12:1–26
Fox-Walsh KL, Dou Y, Lam BJ, Hung SP, Baldi PF, Hertel KJ (2005) The architecture of pre-mRNAs affects mechanisms of splice-site pairing. Proc Natl Acad Sci U S A 102:16176–16181
Frisone P, Pradella D, Di Matteo A, Belloni E, Ghigna C, Paronetto MP (2015) SAM68: signal transduction and RNA metabolism in human cancer. Biomed Res Int
Geuens T, Bouhy D, Timmerman V (2016) The hnRNP family: insights into their role in health and disease. Hum Genet 135:851–867
Glaich O, Leader Y, Lev Maor G, Ast G (2019) Histone H1.5 binds over splice sites in chromatin and regulates alternative splicing. Nucleic Acids Res 47:6145–6159
Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, Krysiak K, Harris CC, Koboldt DC, Larson DE et al (2012) Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat Genet 44:53–57
Grosso AR, Martins S, Carmo-Fonseca M (2008) The emerging role of splicing factors in cancer. EMBO Rep 9:1087–1093
Hall SL, Padgett RA (1994) Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. J Mol Biol 239:357–365
Harbour JW, Roberson EDO, Anbunathan H, Onken MD, Worley LA, Bowcock AM (2013) Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. Nat Genet 45:133–135
Imielinski M, Berger AH, Hammerman PS, Hernandez B, Pugh TJ, Hodis E, Cho J, Suh J, Capelletti M, Sivachenko A et al (2012) Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. Cell 150:1107–1120
Karni R, De Stanchina E, Lowe SW, Sinha R, Mu D, Krainer AR (2007) The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat Struct Mol Biol 14:185–193
König H, Ponta H, Herrlich P (1998) Coupling of signal transduction to alternative pre-mRNA splicing by a composite splice regulator. EMBO J 17:2904–2913
Kornblihtt AR (2005) Promoter usage and alternative splicing. Curr Opin Cell Biol 17:262–268
Kornblihtt AR, Schor IE, Alló M, Dujardin G, Petrillo E, Muñoz MJ (2013) Alternative splicing: a pivotal step between eukaryotic transcription and translation. Nat Rev Mol Cell Biol 14:153–165
Královičová J, Houngninou-Molango S, Krämer A, Vořechovský I (2004) Branch site haplotypes that control alternative splicing. Hum Mol Genet 13:3189–3202
Lev Maor G, Yearim A, Ast G (2015) The alternative role of DNA methylation in splicing regulation. Trends Genet 31:274–280
Li X, Liu S, Zhang L, Issaian A, Hill RC, Espinosa S, Shi S, Cui Y, Kappel K, Das R et al (2019) A unified mechanism for intron and exon definition and back-splicing. Nature 573:375–380

Liu HX, Zhang M, Krainer AR (1998) Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. Genes Dev 12:1998–2012
Long JC, Caceres JF (2009) The SR protein family of splicing factors: master regulators of gene expression. Biochem J 417:15–27
Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T (2010) Regulation of alternative splicing by histone modifications. Science 327(80):996–1000
Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T (2011) Epigenetics in alternative pre-mRNA splicing. Cell 144:16–26
Lynch KW (2007) Regulation of alternative splicing by signal transduction pathways. Adv Exp Med Biol 623:161–174
Maciejewski JP, Padgett RA (2012) Defects in spliceosomal machinery: a new pathway of leukaemogenesis. Br J Haematol 158:165–173
Maguire SL, Leonidou A, Wai P, Marchiò C, Ng CKY, Sapino A, Salomon AV, Reis-Filho JS, Weigelt B, Natrajan RC (2015) SF3B1 mutations constitute a novel therapeutic target in breast cancer. J Pathol 235:571–580
Makishima H, Visconte V, Sakaguchi H, Jankowska AM, Kar SA, Jerez A, Przychodzen B, Bupathi M, Guinta K, Afable MG et al (2012) Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. Blood 119:3203–3210
Manley JL, Krainer AR (2010) A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev 24:1073–1074
Martinez-Contreras R, Cloutier P, Shkreta L, Fisette J, Revil T, Chabot B (2007) hnRNP proteins and splicing control. Adv Exp Med Biol 623:123–147
Matera AG, Wang Z (2014) A day in the life of the spliceosome. Nat Rev Mol Cell Biol 15:108–121
McManus CJ, Graveley BR (2011) RNA structure and the mechanisms of alternative splicing. Curr Opin Genet Dev 21:373–379
Merkin J, Russell C, Chen P, Burge CB (2012) Evolutionary dynamics of gene and isoform regulation in mammalian tissues. Science 338(80):1593–1599
Nilsen TW, Graveley BR (2010) Expansion of the eukaryotic proteome by alternative splicing. Nature 463:457–463
Obeng EA, Chappell RJ, Seiler M, Chen MC, Campagna DR, Schmidt PJ, Schneider RK, Lord AM, Wang L, Gambe RG et al (2016) Physiologic expression of Sf3b1K700E causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell 30:404–417
Padgett RA, Grabowski PJ, Konarska MM, Seiler S, Sharp PA (1986) Splicing of messenger RNA precursors. Annu Rev Biochem 55:1119–1150
Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40:1413–1415
Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, Pellagatti A, Wainscoat JS, Hellstrom-Lindberg E, Gambacorti-Passerini C et al (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med 365:1384–1395
Patel AA, Steitz JA (2003) Splicing double: insights from the second spliceosome. Nat Rev Mol Cell Biol 4:960–970
Pervouchine DD, Khrameeva EE, Pichugina MY, Nikolaienko OV, Gelfand MS, Rubtsov PM, Mironov AA (2012) Evidence for widespread association of mammalian splicing and conserved long-range RNA structures. RNA 18:1–15
Quesada V, Conde L, Villamor N, Ordóñez GR, Jares P, Bassaganyas L, Ramsay AJ, Beà S, Pinyol M, Martínez-Trillos A et al (2012) Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet 44:47–52
Ramanouskaya TV, Grinev VV (2017) The determinants of alternative RNA splicing in human cells. Mol Gen Genomics 292:1175–1195
Reed R (1996) Initial splice-site recognition and pairing during pre-mRNA splicing. Curr Opin Genet Dev 6:215–220

Saint-André V, Batsché E, Rachez C, Muchardt C (2011) Histone H3 lysine 9 trimethylation and HP1γ favor inclusion of alternative exons. Nat Struct Mol Biol 18:337–344
Saldi T, Cortazar MA, Sheridan RM, Bentley DL (2016) Coupling of RNA polymerase II transcription elongation with pre-mRNA splicing. J Mol Biol 428:2623–2635
Scotti MM, Swanson MS (2016) RNA mis-splicing in disease. Nat Rev Genet 17:19–32
Seiler M, Yoshimi A, Darman R, Chan B, Keaney G, Thomas M, Agrawal AA, Caleb B, Csibi A, Sean E et al (2018) H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nat Med 24:497–504
Sharp PA (1994) Split genes and RNA splicing (nobel lecture). Angew Chem Int Ed Eng 33:1229–1240
Sharp PA (2005) The discovery of split genes and RNA splicing. Trends Biochem Sci 30:279–281
Sheth N, Roca X, Hastings ML, Roeder T, Krainer AR, Sachidanandam R (2006) Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res 34:3955–3967
Shin C, Manley JL (2004) Cell signalling and the control of pre-mRNA splicing. Nat Rev Mol Cell Biol 5:727–738
Shirai CL, White BS, Tripathi M, Tapia R, Ley JN, Ndonwi M, Kim S, Shao J, Carver A, Saez B et al (2017) Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. Nat Commun 8
Smith CWJ, Valcárcel J (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci 25:381–388
Staley JP, Guthrie C (1998) Mechanical devices of the spliceosome: motors, clocks, springs, and things. Cell 92:315–326
Stephens PJ, Tarpey PS, Davies H, Van Loo P, Wedge DC, Nik-zainal S, Martin S, Varela I, Bignell GR, Yates LR et al (2012) Stephens-2012-somatic landscape breast ca. 486:400–404
Tripathi V, Sixt KM, Gao S, Xu X, Huang J, Weigert R, Zhou M, Zhang YE (2016) Direct regulation of alternative splicing by SMAD3 through PCBP1 is essential to the tumor-promoting role of TGF-β. Mol Cell 64:549–564
Wahl MC, Will CL, Lührmann R (2009) The spliceosome: design principles of a dynamic RNP machine. Cell 136:701–718
Wan R, Bai R, Shi Y (2019) Molecular choreography of pre-mRNA splicing by the spliceosome. Curr Opin Struct Biol 59:124–133
Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456:470–476
Wang L, Lawrence MS, Wan Y, Stojanov P, Sougnez C, Stevenson K, Werner L, Sivachenko A, DeLuca DS, Zhang L et al (2011) SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N Engl J Med 365:2497–2506
Warf MB, Berglund JA (2010) Role of RNA structure in regulating pre-mRNA splicing. Trends Biochem Sci 35:169–178
Yearim A, Gelfman S, Shayevitch R, Melcer S, Glaich O, Mallm JP, Nissim-Rafinia M, Cohen AHS, Rippe K, Meshorer E et al (2015) HP1 is involved in regulating the global impact of DNA methylation on alternative splicing. Cell Rep 10:1122–1134
Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M et al (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. Nature 478:64–69
Zhang C, Li WH, Krainer AR, Zhang MQ (2008) RNA landscape of evolution for optimal exon and intron discrimination. Proc Natl Acad Sci U S A 105:5797–5802
Zhang J, Zhang YZ, Jiang J, Duan CG (2020) The crosstalk between epigenetic mechanisms and alternative RNA processing regulation. Front Genet 11:1–10
Zhou Q, Derti A, Ruddy D, Rakiec D, Kao I, Lira M, Gibaja V, Chan HM, Yang Y, Min J et al (2015) A chemical genetics approach for the functional assessment of novel cancer genes. Cancer Res 75:1949–1958